AMERICAN JOURNAL of PHYSICS

(Formerly THE AMERICAN PHYSICS TEACHER)

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 9 DECEMBER, 1941 NUMBER 6

Molecular Beams and Nuclear Moments

DONALD R. HAMILTON*

Harvard University, Cambridge, Massachusetts

THE use of *molecular beams*—a general term referring to beams of either atoms or molecules—originated in 1921 with the famous Stern-Gerlach experiment, which gave the first direct evidence of space quantization. In the last ten years the most interesting applications of the molecular beam technic have centered around the measurement of the magnetic moments and angular momentums of atomic nuclei, and have been carried out principally in the laboratory of I. I. Rabi at Columbia University. Out of the methods there developed has grown the new technic of *radiofrequency spectroscopy;* this originated as a means of precision measurement of nuclear magnetic moments but has, in addition, opened up completely new vistas in atomic and molecular physics. The method is most easily thought of as measuring the separation of closely adjacent energy levels of atoms and molecules; the energy differences involved are so small that the frequency of the electromagnetic radiation associated with them—as given by the Bohr frequency condition, $E_m - E_n = h\nu_{mn}$—lies in the radiofrequency range.

Because of the great importance of radiofrequency spectroscopy we shall place on it the primary emphasis of our discussion; and to show in proper perspective its development from the earlier molecular beam experiments, these will also be briefly described.

Perhaps it should be emphasized that the beauty of molecular beam experiments, taken as a whole, lies in the fact that they deal with completely isolated molecules. These molecules are widely separated and suffer no collisions during the experiment—that is, during the time necessary for a molecule to travel the length of the beam as a result of its thermal velocity. Meanwhile they may be subjected to carefully controlled external force fields which will cause their paths to deviate slightly from a straight line. Total deflections as small as a micron (or 10^{-6} of the beam length) produce significant effects. The presence, absence or extent of such deflections are the observables of the experiment, and the interpretation of results in terms of the interaction of a single molecule with the external force field is simple and direct.

1. Molecular beam technic

In this section we shall discuss a few general features common to the various types of molecular beam apparatus. Among such features are (1) a source slit from which molecules effuse, (2) a collimating slit through which pass those molecules whose paths lie within a narrow beam, (3) magnets for deflecting this beam and (4) a means of detecting the molecules of the beam at the end of their journey.

The source slit is usually 0.01 or 0.02 mm wide and a few millimeters in height. Behind it is a chamber in which the substance under inves-

* Society of Fellows, Harvard University.

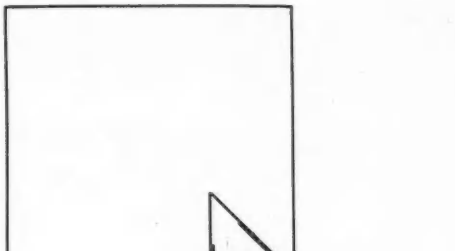


FIG. 1. Two-wire deflecting magnet, cross section perpendicular to beam; field wires enlarged relative to Duralumin block. The limiting positions of the beam are shown by A and B.

tigation exists at a pressure of a few millimeters of mercury. For some substances which have been used, a temperature of 1500°K is necessary to obtain this vapor pressure.

The collimating slit is similar to the source slit. Perhaps contrary to one's first guess, the optimum position is generally midway between source and detector. Both slits may be moved laterally to adjust the position of the beam with respect to the deflecting magnet.

Let us now consider the principles involved in the construction of this deflecting magnet. For this purpose we may visualize the molecule as a top or gyroscope that has a magnetic moment $\boldsymbol{\mu}$ parallel to the axis of rotation. This top consequently experiences a torque when placed in a magnetic field $\mathbf{H}$ and so precesses about $\mathbf{H}$; but there is a net translational force on it only if the field is inhomogeneous. The potential energy of the molecule—or the analogous top—in the field is $-\mu H\cos\theta$, where θ is the angle between $\boldsymbol{\mu}$ and $\mathbf{H}$. Since θ normally remains constant during the precession of $\boldsymbol{\mu}$ about $\mathbf{H}$, the y-component of the net translational force on the molecule is $F_y=\mu\cos\theta\partial H/\partial y$. Clearly this force vanishes when $\mathbf{H}$ is uniform. We have written down the y-component of the force since we shall take the Y-axis as horizontal and perpendicular to the horizontal beam.

Since the beam is high and narrow, we want a deflecting magnetic field in which H varies rapidly in the horizontal, y-direction but is quite constant in the vertical direction. This is accomplished very satisfactorily by employing a "two-wire" deflecting magnet, which consists of two horizontal parallel wires that carry current in opposite directions. Figure 1 shows a cross section perpendicular to the beam. The smaller the radius a of the wires, the larger is the value of $\partial H/\partial y$ for a given value of H. But too much reduction in a makes sufficient water-cooling of the wires impossible and reduces the height and intensity of the beam; the lower limit to a is of the order of 1.5 mm. This limits H roughly to 1200 gauss or less. If stronger fields are required, one may make an iron electromagnet having a field of the same characteristics by using pole faces whose surfaces correspond to equipotentials of the two-wire field. This has the slight disadvantage that H is no longer directly proportional to the exciting current. Figure 2 shows such a deflecting magnet.

Turning now to detection of the beam, we find that the two most extensively used devices are the Stern-Pirani gauge and the Taylor surface ionization detector. The surface ionization gauge

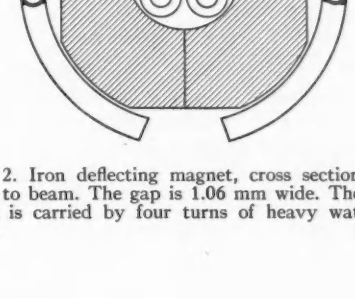


FIG. 2. Iron deflecting magnet, cross section perpendicular to beam. The gap is 1.06 mm wide. The exciting current is carried by four turns of heavy water-cooled copper.

is used for atoms with low ionization potentials —Li, Na, K, Cs, Rb, In, Ga and Ba—or for molecules containing these atoms. Such an atom may be ionized by evaporating it from a metal surface that has a work function higher than the atom's ionization potential, since the surface then attracts the valence electron more strongly than does the atom. In actual practice the beam is allowed to fall on a hot tungsten filament 0.001 or 0.002 in. in diameter, and the beam intensity is measured by the current of positive ions boiled off by this filament. The work function of the pure tungsten filament (4.54 electron-volts) may be raised considerably by playing on it a continuous, minute stream of oxygen; this stratagem makes it possible to use a surface ionization gauge to detect atoms that have ionization potentials slightly above 5.0 electron-volts.

A horizontal cross section of the most recent type of Stern-Pirani gauge is shown in Fig. 3. Four nickel ribbons form the arms of a Wheatstone bridge and are heated by a constant current; their temperature, and hence their resistance, depends upon the rate at which heat is conducted away by the gas which is present. All four ribbons are exposed to the residual gas pressure in the whole vacuum system, but only one pair is exposed to the increased pressure which is built up in one side of the gauge by a beam of molecules of a stable gas. The beam thus unbalances the Wheatstone bridge, and the degree of unbalance is a measure of the beam intensity.

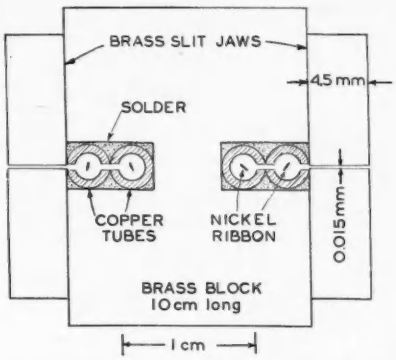


FIG. 3. Horizontal cross section of Stern-Pirani gauge for measuring intensity of beams of stable gases.

2. Atomic hyperfine structure in a magnetic field

The atomic beam experiments which we shall discuss depend on the behavior of the energy levels into which the ground state of an atom is split by the influence of an external magnetic field and a nuclear magnetic moment. Prior to discussing the experiments themselves we shall summarize the general effect of nuclear moments on this behavior. Molecular energy levels are more conveniently discussed in connection with the corresponding experiments (Secs. **6** and **7**).

It is now known that nuclear angular momentums, like electronic angular momentums, have magnitudes of the order of $h/2\pi$. As might be expected from this fact and from the very small charge-mass ratio of nuclei as compared to electrons, nuclear magnetic moments are much smaller than those of electronic states. The latter are of the order of a Bohr magneton μ_0 $[=eh/4\pi mc]$ whereas the former are more conveniently measured in terms of a nuclear magneton, namely, $\mu_0 m/M$, where m and M are the electron and proton masses, respectively.

In the absence of any magnetic fields, either nuclear or external in origin, an atomic state with total extranuclear angular momentum $\mathbf{J}$ will appear as a single energy level.[1] The magnetic field of a small nuclear magnetic moment splits this level, since it introduces an interaction energy of nucleus and electrons, this energy being proportional to $\cos(\mathbf{I}, \mathbf{J})$. The resulting group of closely spaced levels is called a *hyperfine structure multiplet* (hfs multiplet, for short). The total splitting is designated by ΔW when expressed in ergs and by $\Delta\nu$ when measured in the convenient unit "cm^{-1}." ($\Delta W = Lc\Delta\nu$.)

Each level of the hfs multiplet is characterized by a fixed relative orientation of $\mathbf{I}$ and $\mathbf{J}$, that is, by a value of F, the total angular momentum quantum number; F may have the values $I+J$, $I+J-1, \cdots |I-J|$. For a level with given F, $\cos(\mathbf{I}, \mathbf{J})$ is proportional to $F(F+1)$ minus a constant; hence the interval between two adjacent levels F' and $(F'-1)$ is proportional to F'.

[1] $\mathbf{J}$ is a vector giving the angular momentum in units of $h/2\pi$; its magnitude is $[J(J+1)]^{\frac{1}{2}}$, where J is the *angular momentum quantum number*. Similar statements hold for the angular momentum vectors $\mathbf{I}$ and $\mathbf{F}$ and their respective quantum numbers I and F; $\mathbf{I}$ is the *nuclear angular momentum* (or "nuclear spin") and $\mathbf{F}$ is the *total angular momentum* of the atom—that is, $\mathbf{F}=\mathbf{I}+\mathbf{J}$.

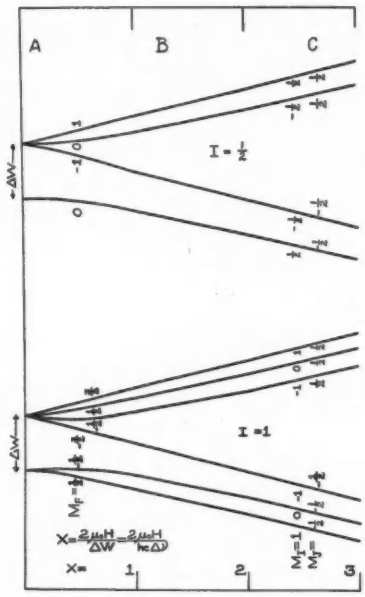


FIG. 4. Hyperfine structure of $^2S_{\frac{1}{2}}$ state for $I=\frac{1}{2}$ (hydrogen) and $I=1$ (deuterium). Each level is labeled with both its weak-field and strong-field quantum numbers. Both diagrams correspond to positive nuclear moments. The weak-field quantum number F is not shown; reading down, the four groups of weak-field levels should be labeled: $F=1, 0, \frac{3}{2}, \frac{1}{2}$.

(This is an example of the familiar Landé interval rule which also occurs in Russell-Saunders coupling in fine structure, where $\mathbf{J}=\mathbf{L}+\mathbf{S}$.)

We shall have to deal only with $^2S_{\frac{1}{2}}$ states, for which there are just two levels, $F=I\pm\frac{1}{2}$. In this case

$$\Delta\nu=\frac{4}{3}\frac{2I+1}{I}\frac{R\alpha^2Z}{n^{*3}}\frac{\mu}{\mu_0}\ \text{cm}^{-1}, \qquad (1)$$

where R is the Rydberg constant, α is the fine structure constant [$=1/137$] and n^* is the effective quantum number.

All this, of course, is for zero external magnetic field; **I** and **J** precess about their resultant, **F**, which remains fixed in direction. In a sufficiently weak, external magnetic field **H**, directed, say, along Oz, **I** and **J** continue to precess rapidly about **F**, which in turn precesses slowly about **H**; **F** may have any one of $2F+1$ orientations characterized by values of M_F (the projection of **F** upon Oz) ranging from F to $-F$ in integral steps.* The additional energy of interaction with **H** is proportional to M_FH. Thus each F-level splits into $2F+1$ equally spaced levels; this is the Zeeman effect of hfs, corresponding to region A in Fig. 4.

When **H** is large enough to be comparable to the field of the nuclear magnetic moment at the position of the electrons, a complicated dynamical situation ensues. For any given state, cos (**I**, **J**) and the magnitude of **F** are no longer constant, although M_F is constant. Even though the quantum number F no longer has a simple physical significance, each level may for convenience be labeled by the values of F and M_F which it had in the weak field.

When H is sufficiently large, **I** and **J** are completely decoupled; they precess independently about **H**, that is, about Oz, their projections on Oz being denoted by M_I and M_J, respectively. This is the Paschen-Back region of hfs, region C in Fig. 4. It corresponds, on a gross scale which overlooks the influence of the nucleus in Fig. 4, to the splitting of the original J state into $2J+1$ levels each characterized by M_J; it is thus the ordinary Zeeman region of fine structure.

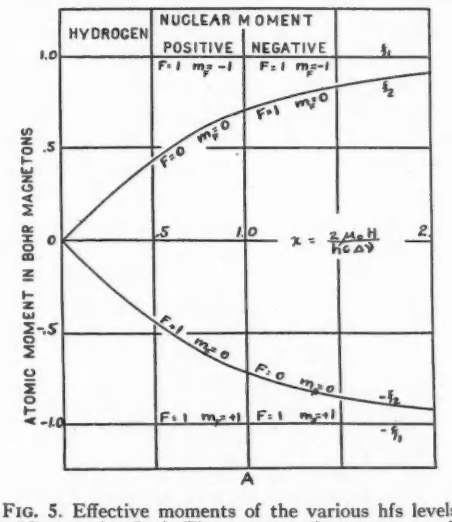


FIG. 5. Effective moments of the various hfs levels of the $^2S_{\frac{1}{2}}$ state for $I=\frac{1}{2}$. The two sets of quantum numbers correspond to positive and negative nuclear moments; the actual proton moment is positive.

* M_F is called a *magnetic quantum number;* more specifically, it is the magnetic quantum number of the total angular momentum. We shall shortly have occasion to refer to M_I and M_J, the magnetic quantum numbers of the nuclear and electronic angular momentums, respectively.

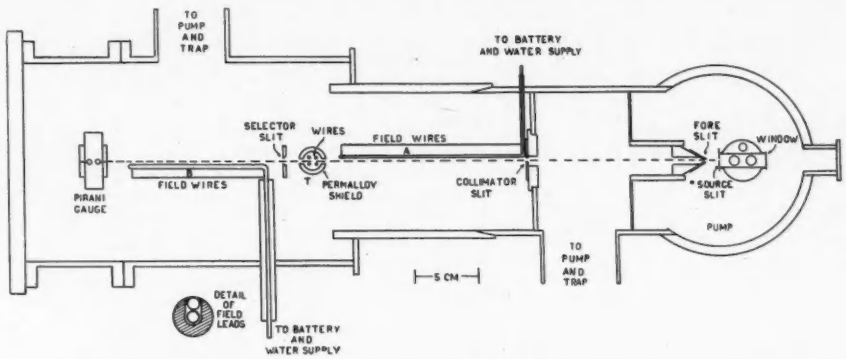


FIG. 6. Apparatus for measuring the magnitudes and signs of proton and deuteron moments.

The energy levels in the Zeeman and Paschen-Back regions may be obtained by the usual vector-model treatment; but the experiments to be described depend greatly on what happens at intermediate fields (region *B*, Fig. 4). The calculation of these energies is a straightforward quantum-mechanical problem which will not be treated here.

3. Early experiments on nuclear moments

We shall now review briefly the earlier molecular and atomic beam experiments on atomic nuclei. In principle, most of the resulting information is obtainable by observation of hfs in atomic spectra; in practice these experiments give information which is either more accurate than that yielded by spectroscopy or is inaccessible to the latter. For example, the experiment which we shall discuss most fully was carried out on hydrogen atoms in the ground state; this state cannot be investigated spectroscopically for hfs because the lines emitted in transitions to it lie in the far ultraviolet region (Lyman series). No other state of hydrogen has a measurable hfs from which to obtain information about the fundamentally important magnetic moments of the proton and deuteron.

The basic quantity in these experiments is the horizontal force exerted on an atom by the deflecting magnet. If E is the energy of the atom in the field $\mathbf{H}$, then

$$F_y = -\partial E/\partial y = -(\partial E/\partial H)(\partial H/\partial y);$$

thus it is the *effective moment* of the atom,

$$\mu_{\text{eff}} = -\partial E/\partial H, \qquad (2)$$

in which we are at present interested. Figure 5 shows μ_{eff} for the various hfs states of hydrogen. Diagrams such as Figs. 4 and 5 are most conveniently plotted in terms of the parameter x, defined for ${}^2S_{\frac{1}{2}}$ states by

$$x = 2\mu_0 H/\Delta W = 2\mu_0 H/hc\Delta\nu. \qquad (3)$$

Since deflection is proportional to effective moment, Fig. 5 exhibits directly the Stern-Gerlach deflection pattern to be expected of a beam of hydrogen atoms with a given velocity. For example, when the value of H is such that $x \sim 0.35$, the four states (which are equally populated) have moments μ_0, $\mu_0/3$, $-\mu_0/3$, $-\mu_0$, respectively; thus the deflection pattern should consist of four equidistant lines of equal intensity. Once the value of H that yields such a pattern has been observed, $\Delta\nu$ can be computed by means of Eq. (3). This example is easily generalized.

A direct deflection experiment based on these principles was carried out in 1934.[2] Since the atoms actually do not all have the same speeds, thus spreading the lines of the deflection pattern into wide overlapping bands, it is necessary, in interpreting the data, to assume the exact velocity distribution of the atoms. The value of the proton moment as calculated from Eqs. (1) and (3) was found to be 3.25 nuclear magnetons ± 10 percent. The discrepancy with the better value of 2.7896 ± 0.0008 nuclear magnetons (known from later experiments) was probably caused by the assumption of an incorrect average velocity of the atoms in the beam.

The risk involved in assumptions about velocities led to the development of another

[2] Rabi, Kellogg and Zacharias, Phys. Rev. **46**, 157 (1934).

method which is independent of such assumptions. The apparatus, shown in Fig. 6, has two two-wire deflecting magnets of different lengths, instead of a single magnet. (The selector slit, field wires T and permalloy shield refer only to the experiment to be described in Sec. 4.) The directions of the currents I_A and I_B are such that their fields $\mathbf{H}_A$ and $\mathbf{H}_B$ are parallel, while $\partial H_A/\partial y$ and $\partial H_B/\partial y$ are opposite in sign. The axis of quantization is the same in fields A and B, and an atom with given value of M_F in field A will have the same value of M_F in field B. Since the gradients are opposite, the deflection caused by A can be exactly compensated by that caused by B, thus giving a net deflection of zero. The whole experiment is based upon this null effect and upon the conditions of its occurrence, which we shall now discuss.

The deflection of an atom in either field is proportional to the corresponding value of $\mu_{\text{eff}}\partial H/\partial y$; thus the condition for zero net deflection is

$$\mu_A\partial H_A/\partial y = K\mu_B\partial H_B/\partial y, \qquad (4)$$

where K is a constant which is calculable from the geometry of the apparatus and which may be shown to be independent of atomic velocity. Since H and $\partial H/\partial y$ are proportional to the current I, we may write $\partial H_A/\partial y = G_A I_A$, and so forth; Eq. (4) then becomes $\mu_A G_A I_A = K\mu_B G_B I_B$. For the states $M_F = \pm 1$, Fig. 5 shows that $\mu_A = \mu_B = \mp\mu_0$, so that I_A/I_B at zero deflection of these states is equal to KG_B/G_A. In the light of this information, observation of I_A/I_B at zero deflection of the states $M_F = 0$ makes it possible to determine μ_A/μ_B, which is a known function of x_A and x_B. But $I_A/I_B = x_A/x_B$; hence we can solve for x_A and x_B. H_A and H_B are calculable from I_A and I_B, respectively, and the geometry of the apparatus; a check on this calculation is provided by comparison of observed and geometrically calculated values of KG_B/G_A. The values of x and H, inserted in Eq. (3), give $\Delta\nu$.

The actual procedure used is to keep I_B constant and observe the intensity at zero deflection as a function of I_A. Two peaks in intensity were observed for hydrogen, in confirmation of the value $I = \frac{1}{2}$. The value of the proton moment was found to be 2.85 ± 0.15 nuclear magnetons; the same experiment, repeated for deuterium, gave 0.85 ± 0.03 nuclear magneton and a spin of 1. Since Eq. (4) is independent of atomic velocity, so are these results.

Basically, these experiments measure $\Delta\nu$ by measuring the value of H required to produce a given degree of decoupling of **I** and **J**; that is, they compare the coupling of **I** and **J** with the known coupling of **J** and **H**. A more direct method, using the direct interaction between **I** and **H**, was employed by Stern and his collaborators.[3] The ground state of the hydrogen molecule has no electronic angular momentum; the only magnetic moments present are those of the protons and the considerably smaller magnetic moment due to rotation of the molecule as a whole. Thus the deflection of H_2 molecules in an inhomogeneous field, together with an assumption about molecular velocities, gives μ_P. The value obtained was 2.46 ± 0.08 nuclear magnetons.

These hydrogen experiments are of far-reaching theoretical importance. If the Dirac equation were applicable to the proton, as it is to the electron, the proton would have a magnetic moment of 1 nuclear magneton; these experiments showed it to have a moment nearly three times as large. This fact is basic in any theories of the fundamental nature of neutron and proton, such as have arisen in connection with the theory of β-decay and of the meson. Furthermore, the substantial agreement between $\Delta\nu$ and the directly measured moment μ_P gave the first really direct confirmation of the origin of hyperfine structure in nuclear magnetic moments. Speculation as to the cause of the slight discrepancy in the two results came to an end when the direct measurements of the proton moment described in Sec. 6 showed that the measurements of Stern's group must have been subject to some unknown error.

A much simpler method of determining $\Delta\nu$ and I is the method of zero moments.[4] This combines the advantages of a velocity-independent null method—measurement of intensity at zero deflection—with the simplicity of a single deflecting magnet. It is based on the circumstance that, for $I > \frac{1}{2}$, there are one or more

[3] Estermann, Simpson and Stern, Phys. Rev. **52**, 535 (1937).

[4] Cohen, Phys. Rev. **46**, 713 (1934).

values of H for which $\mu_{eff}[=-\partial E/\partial H]$ of some of the hfs states passes through zero. For $I=1$, as will be seen from Fig. 4, this occurs only when $x=\frac{1}{3}$; hence in this case, if the beam intensity at zero deflection is observed as a function of H, a single peak will be observed when H has the value corresponding to $x=\frac{1}{3}$.

In general, the number and relative spacing of such peaks determines I; with I known, x is fixed for each peak and observation of the corresponding H's determines $\Delta\nu$ by Eq. (3). This method has been used[5] to observe the $\Delta\nu$'s of the ground states of Na^{23}, Li^{6}, Li^{7}, K^{39}, K^{41}, Cs^{133}, Rb^{85}, Rb^{87}, In^{115}, Ga^{69} and Ga^{71}. In the latter three cases atoms in the metastable state are present in the beam, and a nuclear quadrupole moment perturbs the relative positions of the zero-moment peaks of this metastable state; this fact makes possible an accurate measurement of these quadrupole moments.[6]

The method of zero moments provides a more direct and unambiguous determination of nuclear angular momentums than does spectroscopy, and usually a more accurate measurement of hfs. It generally has the advantage of a larger number of observables than occur in the corresponding spectroscopic case.

4. The sign of the nuclear moment

The experiments of Sec. 3 measure I and $\Delta\nu$. These two quantities, with Eq. (1), give the magnitude of the nuclear magnetic moment but not its sign.[7] To see why this is so we may investigate the effect of a change of sign of the nuclear moment on an hfs pattern.

Reversal of sign of μ reverses the sign of the energy of interaction between electrons and nuclear moment. When $H=0$, this means that the order of the levels in the hfs pattern is reversed. When $H>0$, the energy of interaction between **H** and the electrons enters the picture as the cause of the Zeeman splitting; this energy is independent of the sign of the nuclear moment. Thus, if the low-field region for $I=\frac{1}{2}$ in Fig. 4 were redrawn for negative nuclear moment, $F=0$ would lie above $F=1$, but $F=1$, $M_F=1$ would still lie above $F=1$, $M_F=0$.

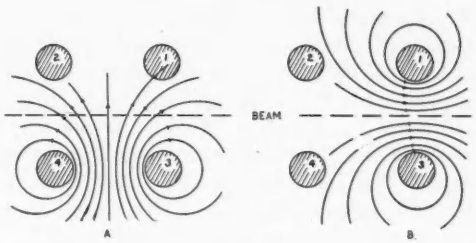


FIG. 7. Magnetic field produced by passing current through different pairs of the field wires T in Fig. 6. A, field for producing transitions in measuring sign of nuclear moment; B, field which produces no transitions.

The same analysis, carried through in general and extended to all values of H, shows that a change in sign of the moment reflects the hfs pattern about a horizontal line and changes the M_F labels on some of the curves. But because of the symmetry of the effective moment diagram there is a change only in its labeling, and none in its form. This is shown, for the particular case of hydrogen, in Fig. 5. Both the zero moment and hydrogen experiments have given information only about the form of the effective-moment diagram; thus we need some additional experimental feature that will differentiate between alternative sets of quantum numbers which label the diagram.

The hydrogen apparatus was designed to provide this additional feature, the auxiliary device being inserted after measuring the magnitude of the moments. It is shown in Fig. 6 as a selector slit and the field wires T. These wires are perpendicular to the plane of the paper, and a small current passed through the proper pair gives the magnetic field shown in Fig. 7A. As "seen" by an atom of the beam, this field is equivalent to a weak magnetic field that rotates through 180° in the time taken for a hydrogen atom to travel between the wires.

Since H_A and H_B are parallel, an atom ordinarily has the same orientation—that is, the same magnetic quantum number M_F—in both fields; the purpose of the new field T is to cause a sudden reorientation—that is, a "nonadiabatic" transition to a state of different M_F—in the region between the deflecting magnets A and

[5] See Millman and Fox, Phys. Rev. **50**, 220 (1936) for bibliography.

[6] Hamilton, Phys. Rev. **56**, 30 (1939); Renzetti, Phys. Rev. **57**, 753 (1940).

[7] A magnetic moment is positive (or negative) if it and the angular momentum are such as to be produced by rotation of a positive (or negative) charge. For example, the electron has a negative moment.

B. This transition-producing field is too weak to cause any change in F. As shown by the exact theory of this effect, the essential condition for transitions is that the angular speed of the field, as "seen" by the atom, be equal to the angular speed of the atom's precession in the field. The field in Fig. 7B should cause no transitions since the field appears to the atom to make no sudden changes in direction.

It will now be recalled that, at one stage in the determination of the $\Delta\nu$ of hydrogen, fields A and B were adjusted so that the atoms in the two states $M_F=0$ had no net deflections at the detector; at the same time, the two states had large and opposite deflections at a point midway between magnets A and B, since the moments of the two states are opposite in sign. With the selector slit one may cut out one of these states, say the one of negative moment, so that only atoms in the state of positive effective moment are allowed to reach the detector. These atoms are now subjected to the transition-producing field. If they are in the state $F=0$, $M_F=0$—that is, if the nuclear moment is positive—then no transitions will occur since there are no other states with $F=0$ to which transitions can be made. But if these atoms are in the state $F=1$, $M_F=0$—negative nuclear moment—transitions may be made to $F=1$, $M_F=\pm 1$. The state $F=1$, $M_F=1$ has negative moment so that, if atoms make the transition to this state, they will in field B be deflected away from the detector instead of being focused in to it. Thus, if the moment is negative, a decrease in intensity will occur whereas a positive moment will give no such result.

The procedure described in the previous paragraph is merely one simple example of the method of nonadiabatic transitions. The various possible experimental combinations all indicate that the proton and deuteron have positive moments.

The same general method has been used to measure the sign of the moments of a number of the alkali metals, all of which were found to have positive moment.[8]

5. Radiofrequency spectroscopy

The experiments described in Sec. 4 are mainly qualitative. For one thing, the necessary equality between the angular speed of **F** about **H** and of **H** about a fixed direction in space cannot be satisfied simultaneously for atoms of two different velocities. Furthermore, H must be less than 1 gauss, hence stray fields make it impossible to know the magnitude and form of the transition-producing field.

These qualitative experiments interest us primarily because of their illustrious offspring, the new type of precision experiment of much broader scope known as radiofrequency spectroscopy. The birth of the latter came with the change of the transition-producing field from a static magnetic field to one that is itself oscillating at a fixed frequency.

This is a radical improvement. Equality between the frequencies of motion of **F** and of **H** may now be fulfilled exactly for all atoms of the beam. Furthermore, since the frequency of motion of **H** is no longer kept small by dependence on atomic velocity, it is no longer necessary to employ a weak field H to keep the frequency of motion of **F** small; and with the magnetic field increased in spatial extent and magnitude, its form and magnitude are much more accurately known.

The magnet which produces this oscillating magnetic field has two parts. One is similar to the magnet of Fig. 2 except that the two pole faces are plane, parallel and vertical, producing a homogeneous static field perpendicular to the beam. The second part is a "hairpin," two wires lying in a horizontal plane between the pole faces of the "homogeneous" magnet; these wires carry, in opposite directions, an oscillating current which produces a weak oscillating magnetic field perpendicular to the homogeneous field. The resultant magnetic field thus oscillates slightly about the fixed field, which is perpendicular to the beam.

Let us now examine the effect of this oscillating field upon a simplified system, one that has a single angular momentum vector **J** and a magnetic moment $\boldsymbol{\mu}$ parallel to **J**. If H_1 (the oscillating component) is quite small compared to H (the fixed component), then **J** precesses about **H** with the Larmor precession frequency,

$$\nu=\mu H/Jh; \tag{5}$$

as usual the orientation of **J** is specified by one

[8] Torrey, Phys. Rev. **51**, 501 (1937); Gorham, Phys. Rev. **53**, 563 (1938)

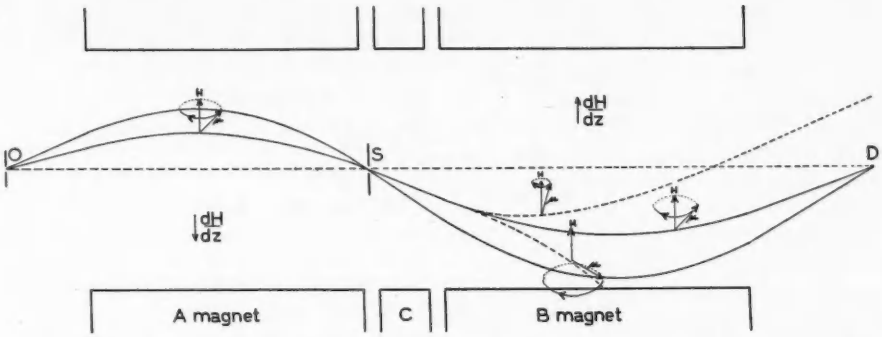


FIG. 8. Paths of molecules. The solid curves are the paths of two molecules of different velocity and moment and whose moments are not changed during passage through the apparatus. The dotted curves indicate paths of two molecules, the projection of whose nuclear magnetic moments along the field has been changed in the region of magnet *C*. This is indicated by the variously oriented gyroscopes drawn on the dotted and solid curves.

of the $(2J+1)$ possible values of M_J. The oscillating component $\mathbf{H}_1$ now subjects $\mathbf{J}$ to an oscillating torque about an axis perpendicular to $\mathbf{H}$.

The effects of the oscillating torque have been worked out quantum-mechanically; what happens is perhaps best understood in terms of an everyday mechanical model. Consider a gyroscope or top precessing about a vertical axis under the influence of gravity and subjected to an oscillatory torque about a horizontal axis, such as might be produced by a sinusoidal displacement of the point of support of the top. During part of a cycle this torque will tend to make the top "stand up" (decrease its angle with the vertical) and at other instants "lie down." The net effect, averaged over a few cycles, is in general nil. But if, and only if, the precession and oscillation frequencies are equal, the effect is cumulative and the top may actually stand up or lie down. The direction of change depends on the phase relation between the oscillating torque and the oscillatory motion of precession.

In the quantum-mechanical language of the analogous atomic system, this phenomenon corresponds to an increase or decrease in M_J—that is, to the occurrence of transitions between states of different M_J. The condition for occurrence of these transitions is that f, the frequency of oscillation of H_1, be equal to the Larmor frequency ν, given by Eq. (5); this condition is suggested by the analogy and is borne out by the quantum-mechanical theory.

A very interesting fact may now be noted. The energy of the system is given by $E=-\mu HM_J/J$; thus Eq. (5) and the condition $\nu=f$ are exactly equivalent to the requirement that f be given by the Bohr frequency condition for electromagnetic radiation emitted or absorbed in transitions between adjacent levels,

$$\Delta E=\mu H/J=hf. \tag{6}$$

It is thus apparent that there are two equivalent ways of looking at the whole problem. One is the dynamical approach, in which the exact details of precession in a specific magnetic field are worked out quantum-mechanically. The other approach considers $\mathbf{H}_1$ as a field of electromagnetic radiation which causes induced emission and absorption of energy in transitions between energy levels of a system, in accordance with the Bohr frequency relations.

Historically, the dynamical approach came first, since the first experiments were carried out on systems whose dynamics are described with sufficient accuracy by our simple model of a system with one angular momentum present. But the corresponding treatment of a system in which several angular momentums interact with one another and with $\mathbf{H}_1$ and $\mathbf{H}$ presents great mathematical difficulties; on the other hand, the calculation of the energy levels of such a system is fairly straightforward. Since the two methods are equivalent, the simplicity of the second makes it preferable for dealing with the question of the frequencies at which transitions occur; from this

viewpoint, this type of experiment lies in the realm of spectroscopy.

The reason for the name *radiofrequency spectroscopy* becomes obvious on consideration of the order of magnitude of frequencies involved. For electronic magnetic moments, $\nu/H \sim \mu_0/h \sim 1.4$ megacycle/gauss; and for nuclei, $\nu/H \sim m\mu_0/Mh \sim 0.75$ kilocycle/gauss.

The change from static to radiofrequency transition-producing field necessitates no essential change in the method of detecting transitions. In the former experiments, deflecting fields and selector slit were so arranged that atoms of only one particular state reached the detector (were "focused") at a time; a more exact control over the transitions, made possible with the new method, makes unnecessary any separate observations of the transitions from each state; hence the apparatus is so arranged that, in the absence of transitions, all molecules will be focused.

Figure 8, a very schematic diagram of a typical apparatus, shows how this is accomplished. The

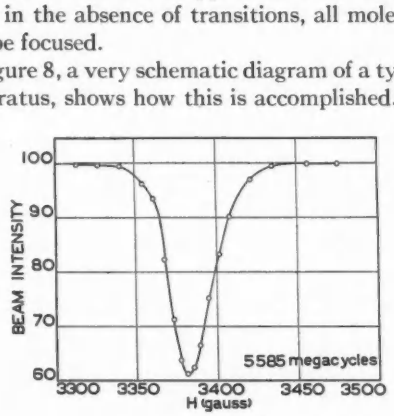


FIG. 9. Resonance curve of the Li^7 nucleus in LiCl.

solid lines are typical paths of molecules in two different states. The slit S is in line with the oven O and detector D. With the deflecting field on, the molecules that pass through S are those that leave O at a slight angle to OS; this angle is different for different states and different velocities. With the gradients of fields A and B in opposite directions, and with the proper adjustments made, all molecules that pass through S will reach D. The intensity of a perfectly focused beam is equal to the intensity with deflecting fields off.

Once a well-focused beam has been obtained, the radiofrequency field is turned on and a search is made for a combination of f and H that causes transitions and thus decreases the intensity of the detected beam. The values of f and H corresponding to these "resonance minimums" are the data of the experiment.

6. The radiofrequency spectroscopy of molecules

The first experiments in radiofrequency spectroscopy[9] (or the *molecular beam magnetic resonance method*, as it is also called) were carried out on molecules—specifically, on molecules that contain an alkali atom for detection purposes and have zero electronic moment so as to make the molecule's moment sensitive to reorientations of the nuclear moment.

With zero electronic moment, the only moments present are those of the nuclei and that due to the rotation of the molecule as a whole. The interactions between these moments are weak; in the magnetic field used in these experiments (a few thousand gauss) each angular momentum precesses practically independently about the external field. This decoupling allows us to apply our simple model of Sec. 5, according to which resonances will occur at the Larmor frequencies of the various nuclei in the molecule. Considerations similar to those of the next section show that the weak interactions which we are neglecting cause a general broadening of a resonance minimum but not a shift of its center.

Figure 9 shows a typical resonance minimum obtained with a beam of LiCl molecules. *A priori*, this might be a resonance due either to the Li or Cl nucleus. The fact that a resonance is found at this same value of f/H when LiF is used, together with the depth of the minimum, shows that the resonance arises from the Li^7 nucleus.

As a check on the theory of the experiment and on the calibration of the apparatus, measurements for every nucleus are made over a wide range of frequencies; the results show excellent consistency. It will be seen from Eq. (5) that the value of f/H for any given nucleus does not yield μ, but only the *nuclear g-factor;* the latter is defined as the ratio of the nuclear moment to the nuclear angular momentum, with the moment measured in nuclear magnetons and the angular

[9] Rabi, Millman, Kusch, and Zacharias, Phys. Rev. 55, 526 (1939).

momentum in units of $h/2\pi$. From Eq. (6), $g=1.312\times10^{-3}f/H$; once I is known, μ may be calculated from g.

The shape of the resonance minimums is significant. The minimum in Fig. 9 is not exactly symmetrical; this is an end effect in the "hairpin." The sign of the asymmetry depends on the sign of the nuclear moment and may be used to determine the latter.[10]

The weakest link in measuring nuclear g-factors is the flip-coil-and-mutual-inductance calibration of magnet C (Fig. 8); in these experiments H is not known to much better than 0.5 percent whereas f is known to 0.005 percent. However, we shall show in Sec. 8 that the radiofrequency spectra of atoms provide a much more accurate calibration of a magnetic field directly in terms of frequency. This method of calibration has recently been used to improve all the g's measured by the method we have just described. The g-values so far observed are listed in Table I.[11]

7. The proton and deuteron moments

The divergence in the values of the fundamentally important proton moment which were obtained from two different types of experiment (Sec. 3) made it important to apply radiofrequency spectroscopy to this question as soon as possible. The apparatus used was similar to that already discussed (Sec. 5 and Fig. 8), except for the use of a source temperature of 78°K instead of the 800°K and more which was necessary for the experiments of Sec. 6. The resulting increase in scope makes these experiments the best example of the power of the new method; we shall therefore discuss them in considerable detail.

TABLE I. *The values of nuclear g-factors observed by the molecular beam magnetic resonance method, the values of nuclear angular momentum, and the resulting values of nuclear magnetic moments. The diamagnetic correction given in the last column should be added to the moments.*

NUCLEUS	OBSERVED g-FACTOR	ANGULAR MOMENTUM I	MOMENT μ	DIAMAGNETIC CORRECTION (%)
$_1H^1$	5.5791±0.0016	1/2	2.7896	0
$_1H^2$	0.8565±0.0004	1	0.8565	0
$_3Li^6$	0.8213±0.0005	1	0.8213	0.01
$_3Li^7$	2.1688±0.0010	3/2	3.2532	0.01
$_4Be^9$	0.784 ±0.003	3/2	−1.176	0.02
$_5B^{10}$	0.598 ±0.003	1	0.598	0.03
$_5B^{11}$	1.791 ±0.005	3/2	2.686	0.03
$_6C^{13}$	1.402 ±0.004	1/2	0.701	0.03
$_7N^{14}$	0.403 ±0.002	1	0.403	0.04
$_7N^{15}$	0.560 ±0.006	1/2	0.280	0.04
$_9F^{19}$	5.250 ±0.005	1/2	2.625	0.06
$_{11}Na^{23}$	1.4765±0.0015	3/2	2.215	0.08
$_{13}Al^{27}$	1.452 ±0.004	5/2	3.630	0.10
$_{17}Cl^{35}$	0.547 ±0.002	5/2	1.368	0.14
$_{17}Cl^{37}$	0.454 ±0.002	5/2	1.136	0.14
$_{19}K^{39}$	0.260 ±0.001	3/2	0.391	0.16
$_{19}K^{41}$	0.143 ±0.001	3/2	0.215	0.16
$_{37}Rb^{85}$	0.536 ±0.003	5/2	1.340	0.39
$_{37}Rb^{87}$	1.822 ±0.006	3/2	2.733	0.39
$_{55}Cs^{133}$	0.731 ±0.002	7/2	2.558	0.67
$_{56}Ba^{135}$	0.558 ±0.002	3/2	0.837	0.68
$_{56}Ba^{137}$	0.624 ±0.002	3/2	0.936	0.68

[10] Millman, Phys. Rev. **55**, 628 (1939).
[11] Millman and Kusch, Phys. Rev. **60**, 91 (1941).

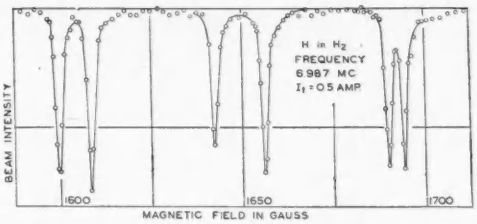


FIG. 10. Resonance curve arising from reorientation of the resultant nuclear angular momentum in H_2.

As was expected, resonances were found with beams of H_2, HD and D_2 molecules;[12] but these resonances, instead of having the simple form of Fig. 9, showed a very pronounced fine structure. The resonance curves associated with reorientation of the proton and deuteron in the three molecules are shown in Figs. 10 to 13. We shall refer to a single minimum in these figures as a *resonance minimum*, and to each fine structure pattern of closely grouped resonance minimums as a *resonance curve*. Similar resonance curves, arising from reorientation of the molecule's rotational angular momentum, are observed for each molecule but are not shown here.

From the positions of the resonance curves as a whole, one may deduce approximate values of the moments involved; but the fine structure indicates that these molecules are not described, with accuracy sufficient for this apparatus, by a simple model which neglects the interaction of the various angular momentums. To jump ahead of our story a little, fine structure is observable because the small moments of inertia and low source temperature allow only several of the lowest rotational states of the molecule to be

[12] Kellogg, Rabi, Ramsey, and Zacharias, Phys. Rev. **56**, 728 (1939).

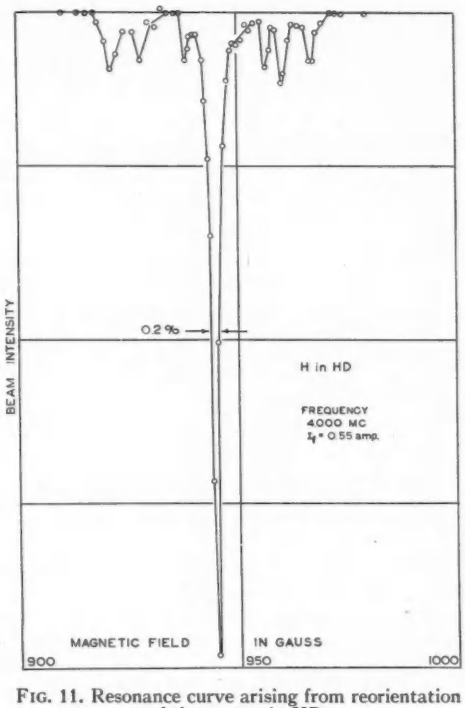


FIG. 11. Resonance curve arising from reorientation of the proton in HD.

excited. If many rotational states were excited, the different fine structures due to each would smear each other out, leaving a single smooth curve centered at a value of f/H corresponding to the nuclear moment. (This is what happened in the experiments of Sec. 6.) To ferret out the information contained in the details of the fine structure we therefore need a comprehensive picture of the relevant energy levels of the molecule, those which arise from a given rotational state and which differ only in the relative orientations of the nuclear and rotational angular momentums. Transitions involving changes in the rotational quantum number J correspond to changes in the energy of the molecule that are too large to be observed. (We use J in this sense in this section only.)

Before proceeding with the theory of these energy levels we pause to note that the data to be obtained from Figs. 10 to 13 are not in orthodox spectroscopic form. That is, this experiment did not measure the various absorption and stimulated emission frequencies of the molecule at a fixed value of the field; instead, it measured the different magnetic fields for which each of these frequencies becomes equal to a certain fixed frequency. This adds no real complications.

Let us now review some of the general features of the molecular states in which we are interested. First of all, in dealing with homonuclear molecules such as H_2 and D_2 we encounter two definite relations that restrict the rotational angular momentum of the molecule and the relative orientation of the two nuclear angular momentums. These relations, familiar from the theory of the specific heats of such molecules, may be described as follows. (1) The nuclear angular momentums act as a single angular momentum vector. Although the allowed values of this vector are given by the usual rules for addition of angular momentums, this is not an ordinary vector coupling arising from an interaction between the two nuclei and so cannot be broken down by applying an external field. This resultant nuclear angular momentum (which we shall in this section denote by **I**) may have the values 0, 1 in H_2 and 0, 1, 2 in D_2. (2) If J is even, I must be even; if J is odd, I must be odd.

It is an empirical fact that only the quantum-mechanical states which obey these restrictions exist in nature, just as it is an empirical fact that, of all possible atomic states, only those which satisfy the Pauli exclusion principle are known to exist. Both the Pauli exclusion principle and the rules which we have just stated are consequences of the particular type of statistical mechanics—or "statistics"—that is applicable to all elementary particles.

On the other hand, these restrictions no longer hold if the two nuclei are different, as in HD. Each nucleus now behaves as an individual; there is no hard and fast resultant nuclear

TABLE II. *Relative abundance of different rotational states at 78°K.*

	H_2			D_2			HD	
J	I	Stat. Wt.	Rel. Abund.	I	Stat. Wt.	Rel. Abund.	Stat. Wt.	Rel. Abund.
0	0	1	0.248	0,2	6	0.559	6	0.628
1	1	9	.745	1	9	.328	18	.369
2	0	5	.003	0,2	30	.105	30	.003
3	1	21	.000	1	21	.004	42	.000

angular momentum quantum number and no relation between nuclear orientation and the allowed value of J.

Table II gives the abundances of the relevant allowed states at the source temperature of 78°K (liquid nitrogen temperature). In a magnetic field each of these states splits up into $(2I+1)(2J+1)$ levels for H_2 or D_2 or $6(2J+1)$ for HD. These numbers are the statistical weights given in Table II. The magnetic fields used are strong enough to decouple **I** and **J**, so that these levels are best labeled with values of M_I and M_J, the magnetic quantum numbers of the resultant nuclear and the rotational angular momentum, respectively.

Consider first the molecules with $J=0$. The H_2 molecules make no transitions since there is only one state ($I=0$). Some of the D_2 molecules are likewise in this inert state, $I=0$, but most of them are in the state $I=2$ for which $\mu=2\mu_D$. Here we have an example of a system with only one angular momentum vector, like the simplified model discussed in Sec. 6. This system has five energy levels spaced $\Delta E=2\mu_D H/2=\mu_D H$ apart, and hence gives rise to a single resonance minimum at the Larmor frequency of the deuteron, $\nu=\mu_D H/h$. Consideration of the relative abundances in Table II shows that this

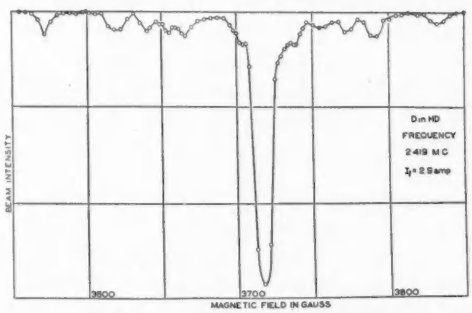


FIG. 12. Resonance curve arising from the reorientation of the deuteron in HD.

resonance minimum must correspond to the deep central minimum in Fig. 13 and that the smaller minimums probably arise from $J=1$.

Although the state $J=0$ of HD has no fixed resultant nuclear angular momentum, the situation is almost as simple as in the cases of H_2 and D_2. In a diatomic molecule, the internuclear axis is perpendicular to **J**; but, when $J=0$, this

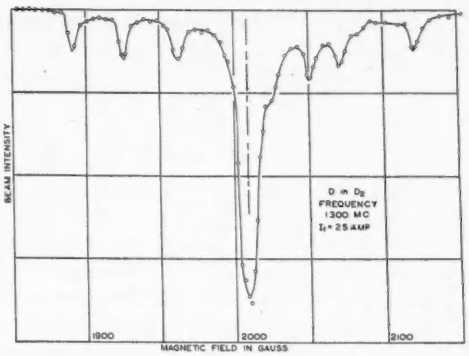


FIG. 13. Resonance curve arising from the reorientation of the resultant nuclear angular momentum in D_2.

axis may have any orientation with respect to **H** or the nuclear angular momentums. But the energy of interaction of two magnetic dipoles—the nuclei—averaged over all directions of the line joining them, is zero; hence the two nuclei may rigorously be considered to precess independently about **H**. If M_P and M_D are the respective magnetic quantum numbers, the energy levels are given by $E=-2\mu_P M_P H - \mu_D M_D H$, and single resonances will be observed at the Larmor frequencies of the proton and deuteron. These resonances correspond to the deep central minimums in Figs. 11 and 12.

Thus the resonances observed for $J=0$ suffice in themselves to determine μ_P and μ_D. The values obtained are $\mu_P=2.785\pm0.02$ and $\mu_D=0.855\pm0.006$ nuclear magneton. The uncertainties arise principally from the calibration of magnet C. The ratio μ_P/μ_D, measured in such a way as to make it independent of this calibration, was found to be 3.2570 ± 0.001.

For $J=1$, the situation is more complicated. The internuclear axis is now limited in its orientation with respect to the external field so that the energy of one nucleus in this external field depends upon the orientations of the other nucleus and of **J**. Thus in H_2 ($J=1$ and $I=1$) we will expect nine energy levels, corresponding to three possible orientations for **I** and three for **J**. In strong fields the energies will be almost but not quite given by $E=-2\mu_P M_I H - \mu_R M_J H$. We say "not quite," since this equation neglects the interactions of the nuclei with each other and with the magnetic field set up by the molecular rotation.

TABLE III. *Magnetic fields at which resonances will occur for a fixed oscillator frequency* $f_0=2\mu_P H_0/h$.

M_J	ΔM_I	MAGNETIC FIELDS
1	0→1	$H_0+(3H''/5-H')\quad -(3H''/5-H')^2/H(1-\alpha)$
1	−1→0	$H_0-(3H''/5+H')\quad -6H''/5(3H''/5+2H')/H(1-\alpha)$
0	0→1	$H_0-6H''/5\quad -(3H''/5-H')^2/H(1-\alpha)$
0	−1→0	$H_0+6H''/5\quad -(3H''/5-H')^2/H(1-\alpha)$
−1	0→1	$H_0+(3H''/5+H')\quad -6H''/5(3H''/5+2H')/H(1-\alpha)$
−1	−1→0	$H_0-(3H''/5-H')\quad -(3H''/5-H')^2/H(1-\alpha)$

The selection rules, $\Delta M_I=0$, $\Delta M_J=\pm1$ and $\Delta M_I=\pm1$, $\Delta M_J=0$, govern transitions between these energy levels; each of these two pairs of selection rules corresponds to six transitions. The group for which $\Delta M_I=\pm1$ corresponds to the six resonance minimums shown in Fig. 10; a similar resonance curve arises from the transitions $\Delta M_J=\pm1$ but is not shown here.

One may hope to understand the details of Fig. 10 when the various intramolecular interactions are taken into account. As a tentative energy expression which includes all the obvious interactions, we take

$$E=-2\mu_P\mathbf{I}\cdot\mathbf{H}-\mu_R\mathbf{J}\cdot\mathbf{H}-2\mu_P H'\mathbf{I}\cdot\mathbf{J}+\langle\mu_P^2/r^3\rangle_{Av}\{\mathbf{i}_1\cdot\mathbf{i}_2-3(\mathbf{i}_1\cdot\mathbf{r})(\mathbf{i}_2\cdot\mathbf{r})/r^2\}. \quad (7)$$

The first and second terms represent the interactions between nuclear and rotational moments and external field. The third term gives the energy of the nuclei in the magnetic field $\mathbf{H}'$ produced at the position of the nuclei by the rotation of the molecule. (This field is parallel to $\mathbf{J}$.) The last term represents the interaction of the magnetic moments of two protons, which have angular momentums $\mathbf{i}_1$ and $\mathbf{i}_2$ and are separated by the internuclear distance $\mathbf{r}$.

From Eq. (7) one may calculate quantum-mechanically the exact energies of the molecular states in the magnetic field; the energy differences and the use of some simple algebra then give the magnetic field at which the frequency for each transition will equal any particular fixed frequency f_0. The results are given in Table III, where

$$H_0\equiv hf_0/2\mu_P, \quad H''\equiv\langle\mu_P/r^3\rangle_{Av}, \quad \alpha\equiv\mu_R/2\mu_P. \quad (8)$$

Thus H'' is the field produced at the position of one proton by the other proton, when $\mathbf{i}_1$ and $\mathbf{i}_2$ are parallel to $\mathbf{r}$.

From Table III it is evident that the resonance curve should be symmetrical except for the influence of terms that involve H'^2/H, H''^2/H and $H'H''/H$; and, when these terms are neglected, the positions of the six minimums are determined by the parameters H' and H''. Furthermore, the center of symmetry should be the field H_0 given by Eq. (8).

Thus one would expect the slightly asymmetrical pattern in Fig. 10 to become symmetrical for larger values of H. This is found to be so; and the observed dependence of this asymmetry on H is sufficient to determine which transition gives rise to a particular minimum in Fig. 10. And now the crucial test of Eq. (7) consists of its ability to fit the positions of the six minimums with consistent values of only two parameters, H' and H''. It is found that the resonance curves are perfectly accounted for by the values

$$H'=27.2\pm0.3 \text{ gauss}, \quad H''=34.1\pm0.3 \text{ gauss}.$$

Furthermore, the value of μ_P calculated from the position of the center of the pattern agrees with that previously obtained from the states $J=0$.

We now recall that $H''=\langle\mu_P/r^3\rangle_{Av}$. But $\langle1/r^3\rangle_{Av}$ is known from band-spectroscopic data to be 2.438×10^{-24} cm^3; hence, from the value of H'' we may calculate μ_P, and we find it to be 2.785×0.03 nuclear magnetons. This determination of μ_P from the interaction of the protons with each other is independent of the f/H determination based on the interaction of a proton and an external field; and, as a matter of fact, an error in calibration of the magnetic field would affect the two values in opposite directions. These two independent determinations of μ_P are in almost too perfect agreement.

The complete consistency of all these results gives the strongest support to the validity and sufficiency of Eq. (7).

Next consider the state of D_2 with rotational quantum number $J=1$; here, just as in H_2,

$I=1$. Thus the discussion of the energy levels of H_2 for $J=1$ should hold without change for D_2. The expected values of H' and H'' are readily calculable from those already obtained for H_2. A little consideration of angular velocities shows that H' should be half as large as in H_2; similarly, $H''[\equiv\langle\mu_D/r^3\rangle_{Av}]$ is already known. Thus the expected values are $H'=13.6$ gauss and $H''=10.5$ gauss. This corresponds to a pattern of six lines with a total width of about 40 gauss.

This is not at all what is observed; the pattern in Fig. 13 has six lines but a total width of about 230 gauss. This great discrepancy is all the more astonishing in view of the complete adequacy of Eq. (7) for H_2.

The logical place to look for some more light on this situation is in the resonance curves of HD, Figs. 11 and 12. The curve that results when the proton is reoriented and the deuteron remains undisturbed is shown in Fig. 11; the pattern is about 50 gauss wide and exactly what would be expected from the results with H_2 and the interactions of Eq. (7). But the reorientation of the deuteron in HD gives a resonance curve (Fig. 12) about 265 gauss wide, which is just as unorthodox as the 230-gauss resonance curve of the deuteron in D_2.

This indicates pretty definitely that there is some unforeseen interaction between the deuteron and the rest of the molecule which is not taken into account in Eq. (7). To make a long story[13] short, a satisfactory solution to this quandary has been found in the hypothesis that, pictorially speaking, the charge distribution of the deuteron is slightly elongated in the direction of the deuteron's angular momentum axis. At the deuteron the electric field set up by the rest of the molecule—nucleus and electrons—is quite nonuniform; and the electrostatic energy of an asymmetrical deuteron in this field would depend on the deuteron's orientation with respect to the field, thus providing the missing interaction.

Such asymmetry of charge density is known to exist in many heavier nuclei, where it affects the spacing of atomic hfs levels. The asymmetry is measured by Q, the electric quadrupole moment, which is defined by the equation

$$eQ=e\int\rho(3z^2-r^2)\mathrm{d}\tau, \qquad (9)$$

where $e\int\rho\mathrm{d}\tau=Ze$, and the integration is carried out for the state $M_I=I$.

Thus we add to Eq. (7) the proper term involving Q and see if the modified equation is now capable of describing the data. We shall not give the details of the analysis, since it is similar to that already gone through for H_2. Suffice it to say that, in addition to H' and H'', there is now available a third parameter H''', defined by $H'''=-5e^2qQ/4\mu_D$. Here q is defined by an equation similar to Eq. (9) but involving an average over all the charges in the molecule except the deuteron; q is a measure of the nonuniformity of the electric field.

The modified energy level equation fits the individual resonance curves perfectly, and the values of H''' calculated from different resonances are in excellent agreement. The reorientation of D in D_2 gives $H'''=87.5\pm1$ gauss; of D in HD, $H'''=(87.5\pm1)$ gauss. The average of these values and of those obtained from the reorientation of **J** in HD and D_2 is $H'''=(87.2\pm0.5)$ gauss. On calculating q from the wave functions of the hydrogen molecule, we find the final value for the electric quadrupole moment of the deuteron to be

$$Q=(2.73\pm0.05)\times10^{-27}\ \mathrm{cm}^2.$$

Pictorially, this is the quadrupole moment of a prolate spheroid deuteron; the major axis, about which the deuteron "spins," is ~15 percent longer than the minor axis.

We have mentioned resonance curves arising from reorientations of **J**. To complete the picture of the information obtained in this series of experiments we state some additional results obtained from these resonance curves.[14] The respective rotational magnetic moments of H_2, HD and D_2 when $J=1$ are found to be 0.8787 ± 0.0070, 0.6601 ± 0.0050 and 0.4406 ± 0.0030 nuclear magneton; these are in the ratio 4 : 3 : 2. The state $J=2$ of H_2 is found to have a moment exactly twice that of the state $J=1$. Furthermore, the data yield a value for the diamagnetic susceptibility and its variation with the orientation of the molecule. All of this information is either brand new or has been obtained for the first time in an exact form.

In looking over the varied results of these experiments on the hydrogens, one is astonished by the range of information obtained from what started out as a simple measurement of the proton and deuteron moments. The discovery of

[13] Kellogg, Rabi, Ramsey, and Zacharias, Phys. Rev. 57, 677 (1940).

[14] Ramsey, Phys. Rev. 58, 226 (1940).

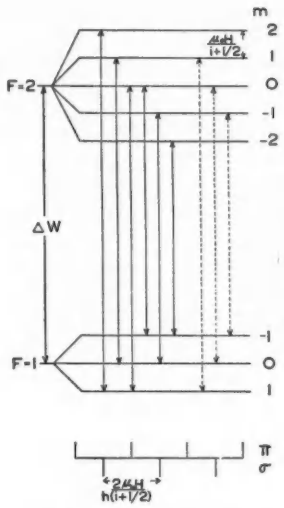


FIG. 14. Low-field hfs magnetic levels for $^2S_{\frac{1}{2}}$ state with $I=3/2$, and allowed transitions between these levels. The lower part of the figure indicates the Zeeman pattern resulting from these transitions.

the deuteron's quadrupole moment is the most exciting of these results; but the molecular data are very valuable to the student of molecular structure.

The values of the deuteron's magnetic moment and electric quadrupole moment affect very fundamentally the theories of the forces between nuclear particles, since the deuteron is the simplest compound nucleus and the most susceptible to exact theoretical treatment. When these experiments were begun, the neutron-proton forces were thought to be central forces, and as a result, the deuteron was considered to be in a $^3S_1$ state, which has zero orbital angular momentum and cannot have any charge asymmetry. The existence of the quadrupole moment indicates the existence of noncentral proton-neutron forces and supports the meson theory of nuclear forces.

Closely associated with the quadrupole moment is the relation of the deuteron magnetic moment to the moments of the proton and the neutron. The neutron moment is (-1.935 ± 0.02) nuclear magnetons, as measured by a radio-frequency technic;[15] hence, within the experimental error, the deuteron moment is the sum of the proton and neutron moments. Slight non-additivity is expected because of the presence of orbital angular momentum as indicated by the quadrupole moment; but the moment of the neutron is not yet known with accuracy sufficient to check this prediction.

[15] Alvarez and Bloch, Phys. Rev. **57**, 111 (1940).

8. The radiofrequency spectroscopy of atoms[16]

It was only natural that these truly spectroscopic experiments on the hydrogens should be followed by a return to the field of atomic hfs with which the earlier atomic beam experiments had been concerned. As we shall see, hfs is still a very fruitful field for research although it is no longer the primary source of information about nuclear moments.

The energy differences involved in hfs are capable of direct observation by radiofrequency spectroscopy. Most hfs splittings of ground states lie between 0.005 and 0.4 cm^{-1}; the range of wave-lengths of electromagnetic radiation associated with transitions between such energy levels is 200 to 2.5 cm; the frequency range is 150 to 12,000 megacycle/sec, the upper end of the technical ultra-high frequency radio range.

In ordinary spectroscopy, hfs is a second-order effect; here it is a first-order effect involving absorption and stimulated emission of quanta in transitions directly between the hfs levels themselves. (These transitions presumably also occur spontaneously in nature; but they occur so slowly that direct observation would be extremely difficult, to say the least.) This makes it possible to measure the position of hfs levels to better than 10^{-6} cm^{-1}, as compared with 10^{-3} cm^{-1} by previous means; a striking example of this high resolving power is provided by the accurate observation of the Zeeman splitting of an atomic energy level in a magnetic field of 0.05 gauss.

The change from beams of molecules to beams of atoms does not involve any essential change in apparatus. Weaker magnetic fields are used because the over-all moments of most atoms are of the order of Bohr magnetons rather than nuclear magnetons; a magnet that will deflect a molecule only slightly will sweep atoms completely out of the beam. And, in the experiments to be described, it was found possible to take readings of beam intensity when the field is constant and the frequency is variable; this makes the data spectroscopically orthodox and somewhat easier to interpret.

[16] Kusch, Millman and Rabi, Phys. Rev. **57**, 765 (1940); Millman and Kusch, Phys. Rev. **58**, 438 (1940).

The goal of these experiments is the measurement of $\Delta\nu$ and I. The most obvious way to do this is to observe the transitions $\Delta F=\pm 1$ in the Zeeman region. The selection rule is $\Delta M_F=\pm 1,0$ and the frequencies exhibit the Zeeman effect of the lines which would be present if the transitions took place in zero field. The center of symmetry of this Zeeman pattern gives directly the energy difference of the two levels in zero field. When $J=\frac{1}{2}$ (which is the case in all the experiments discussed here), there are only two zero field levels (Fig. 4) and their separation is $\Delta\nu$. It is also obvious that such a Zeeman pattern reveals the values of F in the initial and final levels, and hence the value of I.

Let us consider as an illustrative example the hfs of the $^2S_{\frac{1}{2}}$ ground state of K^{39}, for which $F=1,2$ since $I=3/2$. The levels into which the states $F=1,2$ split in a weak magnetic field, the allowed transitions between these levels and the resulting Zeeman pattern are indicated schematically in Fig. 14. The transitions $\Delta M_F=\pm 1$ (solid lines, π-components) arise from the component of the oscillating field perpendicular to the constant field, while the σ-lines arise from the parallel component.

Figure 15 shows a typical experimentally observed pattern for this case. This well-resolved pattern is the Zeeman effect of the line $\Delta F=\pm 1$ for K^{39}, in the very weak magnetic field of 0.25 gauss. A comparison with Fig. 14 identifies the four deepest minimums as π-components; only two σ-components appear, and these are less intense than the π-components. The missing σ-component corresponds to the transition $F=2$, $M_F=-1 \leftrightarrow F=1$, $M_F=-1$; for the field of the deflecting magnets used in this particular instance ($\sim$80 gauss) the moments of initial and final states are closely the same, so that this transition is not detected even though it does occur. The σ-components are of course the weaker since the oscillating field is practically perpendicular to the fixed field; as the oscillating field is decreased in strength, the σ-components may be made to disappear entirely.

From a number of curves similar to Fig. 15, $\Delta\nu$ for the ground state of K^{39} is found to be (461.75 ± 0.02) megacycle/sec, expressed directly in frequency units, or 0.015403 cm^{-1}. This result is independent of the magnetic field in which the transitions occur, since a sufficiently weak field affects only the spacing of the pattern and not its center of symmetry.

However, many $\Delta\nu$'s are larger than are the convenient operating frequencies; the maximum frequency of the oscillator used in these experiments was 700 megacycle/sec, whereas, for Cs^{133}, $\Delta\nu\sim 9300$ megacycle/sec. To measure large $\Delta\nu$'s one must have recourse to observation of intermediate and Paschen-Back spectra, where some of the frequencies are much lower than $\Delta\nu$. Quantum mechanics gives the positions of the energy levels as a function of H, in terms of I, $\Delta\nu$ and the nuclear moment μ. Conversely, the quantum mechanics of the energy levels allows one to calculate $\Delta\nu$, I and μ from the observed spectra.

The moment μ is here a directly observable quantity because of a fine structure in the strong-field spectra which arises from the direct interaction of $\boldsymbol{\mu}$ and $\mathbf{H}$; this interaction is negligible when the field is weak. As to the frequencies in the strong-field spectra, those for the Paschen-Back transitions ($\Delta M_J=0$, $\Delta M_I=\pm 1$) approach $\Delta\nu/(2I+1)$ as H increases. If a still greater reduction in the working frequency is desired, one may resort to the transitions $\Delta F=0$, $\Delta M_F=\pm 1$ in the intermediate region; these frequencies approach zero as H approaches zero.

The calculation of $\Delta\nu$, μ and I from an intermediate or Paschen-Back spectrum involves the assumption that the quantum-mechanical expression for the energy levels is accurate; this

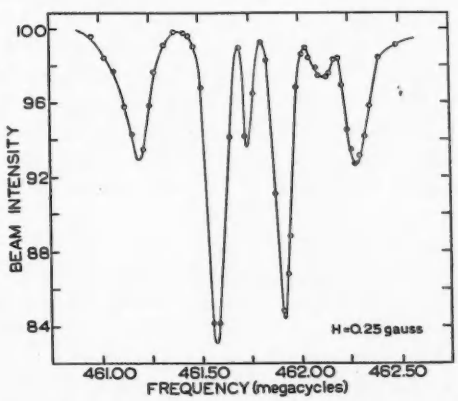


FIG. 15. The Zeeman pattern of the line $\Delta F=\pm 1$ of K^{39}.

accuracy may be tested experimentally by comparing a known μ and a $\Delta\nu$ observed in a Zeeman spectrum with the $\Delta\nu$ and μ calculated from the Paschen-Back and intermediate spectra of the same atom. This comparison has been made with the spectra of K^{39} and Li^7; the perfect agreement which is obtained indicates that the quantum-mechanical calculation is as accurate as the experimental results.

The numerical results of all these experiments on the hfs of various atoms are summarized in Table IV.

We have already indicated that the calculation of $\Delta\nu$ from the Zeeman spectrum is independent of any knowledge of the magnetic field. Without going into details, we may say that the same is true of the Paschen-Back spectrum and practically true of the intermediate spectrum. At the same time, the frequencies of some individual lines in these spectra are sharply dependent on H. Hence, if $\Delta\nu$ is known from the spectrum as a whole, H is known with nearly equal accuracy from some single line. This means of calibrating a magnetic field has been used to increase the accuracy of the experiments of Sec. 6 and is the crucial factor in the high precision of the results given in Table I.

To return to hfs, the values of $\Delta\nu$ obtained for Li^6 and Li^7 are of particular interest. From Eq. (1) and the experimental data we have

$$\frac{\mu_7}{\mu_6}=\frac{(I/2I+1)_7(\Delta\nu)_7}{(I/2I+1)_6(\Delta\nu)_6}=3.9610\pm0.0004. \quad (10)$$

Values of μ_7 or μ_6 calculated from the $\Delta\nu$'s are unreliable since Eq. (1), even in a more accurate form, involves inexact electron wave functions; but the accuracy of the foregoing ratio should be limited only by the accuracy of the $\Delta\nu$'s, since the wave functions are the same for both isotopes. This assumes that the hfs of any $^2S_{\frac{1}{2}}$ state is due solely to the magnetic interaction of a nuclear magnetic dipole moment and the electrons. Conceivably there might be some other factor contributing slightly to hfs, such as a spin-dependent nonelectromagnetic force. One would expect such a coupling to affect the hfs of Li^6 and Li^7 quite differently, since those two nuclei have quite different structures; and, as a consequence, the value of μ_7/μ_6 given by Eq. (9) would be in error.

The true value of μ_7/μ_6 is known from the molecular experiments described in Sec. 6. This true value (measured with the field constant and the frequency variable, to make the ratio of the moments independent of magnetic field calibration) is 3.9610 ± 0.0015, in perfect agreement with the value from hfs. Thus, within the accuracy of these experiments, there is no leeway for additional interactions unless they are very similar in form to the magnetic interaction.

The great precision of the radiofrequency measurement of hfs also opens up the possibility of measuring the quadrupole moments of light nuclei by observing the deviations from the interval rule caused by these moments; such deviations, if they exist, are so small for light atoms as to have escaped detection by the usual spectroscopic means.

Discussion

The importance of precision measurements of nuclear moments is not something to be taken only on faith, in the hope of making spectacular discoveries such as that of the deuteron quadrupole moment. These measurements answer a crying need in the present state of nuclear physics. We have already spoken of the bearing of the hydrogen experiments on the choice of nuclear forces (Sec. 7); the magnetic moments of other light nuclei provide a very sensitive criterion of nuclear models. Nuclear forces and nuclear models are two of the most prominent questions in nuclear theory.

It should be emphasized that a new source of information about molecular structure has been developed. Aside from the hydrogen experiments,

Table IV. *Hyperfine structure $\Delta\nu$'s of ground states of atoms, as measured by radiofrequency spectroscopy.*

Atom	$\Delta\nu$ (megacycle/sec)	$\Delta\nu$ (cm^{-1})
Li^6	228.22±0.01	0.007613
Li^7	803.54±0.04	.026805
Na^{23}	1771.75±0.07	.059103
K^{39}	461.75±0.02	.015403
K^{40}	1285.65±0.1	.042887
K^{41}	254.02±0.02	.008474
Rb^{85}	3035.7 ±0.2	.10127
Rb^{87}	6834.1 ±1.0	.22797
Cs^{133}	9191.4 ±0.9	.30661

there are some unpublished—and not entirely explained—molecular effects occurring in the measurement of nuclear moments in other molecules. It seems fair to say that further molecular developments will go hand in hand with increasing precision in the nuclear measurements. Most molecular effects will probably turn up as by-products of the nuclear studies, since, in general, only the states involved in nuclear reorientations are spaced closely enough to fall within the range of radiofrequency spectroscopy.

Turning to an entirely different aspect of these experiments, it is interesting to note that they have extended the range of application of the quantum theory into the hitherto purely classical region of radiofrequency electromagnetic radiation; and, whereas the frequency associated with a quantum of radiant energy has always hitherto been measured in terms of a length λ and a velocity c, or in some more complicated manner, the frequencies of these radiofrequency quanta are measured directly in terms of a time, $1/f$. This is the period of a quartz crystal oscillator, which in turn may be referred directly to astronomical time.

As to the future of radiofrequency spectroscopy, only a beginning has been made on the magnetic moments, electric quadrupole moments and angular momentums of stable nuclei. The principal problem to be surmounted here is the development of technics for the detection of beams of the relevant molecules and atoms.

More exciting is the imminent prospect of measuring angular momentums and magnetic moments of radioactive nuclei, and correlating these data with information on the β- and γ-ray spectra of these nuclei. An auspicious beginning has been made by Zacharias in his ingenious measurement[17] of the angular momentum and moment of K^{40}, an experiment which we have not described. This isotope makes up about 0.01 percent of natural potassium and is β-active with a period of about 10^8 yrs. The theory of β-decay interprets this long lifetime as due to a large change in I in the decay process; a value $\Delta I \geqslant 3$ was found by calculation to be necessary. The experiment showed that, for K^{40}, $I=4$; K^{40} decays to Ca^{40}, which has a probable nuclear angular momentum of zero so that the predictions of β-theory in this respect are verified.

[17] Zacharias, Phys. Rev. **60**, 158 (1941).

The Origins and Developments of the Concepts of Inductance, Skin Effect and Proximity Effect

THOMAS JAMES HIGGINS
Tulane University, New Orleans, Louisiana

OERSTED AND AMPÈRE

THE systematic investigation of electromagnetic phenomena began in the winter of 1819–1820, when Oersted[1] discovered that a compass needle placed in the vicinity of a circuit traversed by a current is deflected in the same manner as when placed in the field of a permanent magnet. Some months later, in a paper presented before the Académie des Sciences, Ampère[2] described the series of experiments wherein he investigated and from which he formulated the laws governing the mutual mechanical forces exerted between current-carrying conductors. Shortly thereafter, Biot and Savart[3] found that, in an electromagnetic field produced by current in a straight, infinitely long filament, the magnetic field intensity at any

[1] J. C. Oersted, "Experiments on the effect of a current of electricity on the magnetic needle." Annals Phil. [1] 16, 273–276 (1820).

[2] M. A. Ampère, "Memoire sur l'action mutuelle entre deux courans électriques, entre un courant électrique et un aimant ou le globe terrestre, et entre deux aimants," Ann. de Chem. et Physique **15**, 59–76, 170–218 (1820).

[3] J. Biot and F. Savart, "Note sur le magnetism de la pile de Volta," Ann. de Chem. et Physique **15**, 222–223 (1820): ". . . Conduits au résultat suivant qui exprime rigoureusement l'action éprouvee par une molécule de magnétisme austral ou boréal placée à une distance quelconque d'un fil cylindrique trés-fin et indéfini, rendu magnétique par le courant voltaique. Par le point ou réside cette molécule, menez une perpendiculaire à l'axe de fil: la force qui sollicite la molécule est perpendiculaire à cette ligne et l'axe de fil. Son intensité est réciproque à la simple distance."

point a radial distance r from the conductor is, in magnitude, proportional to the reciprocal of this distance r, and is, in direction, perpendicular to the radius vector drawn from the filament to the field-point in question. From this observation the so-called Biot-Savart law can be obtained with but little difficulty. During the next five years Ampère continued his experiments in electrodynamics. In 1825 he advanced the result of these years of labor: an inclusive mathematical theory of electrodynamics, based on his own experiments.[4] A half-decade later Faraday discovered electromagnetic induction.

Michael Faraday

In the spring of 1813 Faraday was appointed to an assistantship in the laboratory of the Royal Institute. Successively assistant to Davy in the latter's chemical researches, director ofthe chemical laboratory and occupant of the chair of chemistry created for him in 1833, his association with the Institute encompassed the whole of his active scientific career, a span of approximately 42 years. Although his initial researches were in chemistry, his paper[5] of 1821 on the history of electromagnetism—in which he reviewed the work of previous investigators, having first duplicated the experiments of those whom he discussed—revealed an active interest in electrophysics, an interest culminating ten years later when, in a paper[6] read before the Royal Society, he summarized the results of a series of notable and important experiments, the fruits of the interim.

The incentive of these experiments was to determine whether or not electromagnetic analogies to certain electrostatic phenomena were to be found.[7] An electrostatic charge causes a charge of opposite sign to appear on adjacent portions of neighboring conductors; if a current were set up in a circuit, might not an adjacent circuit be traversed by an induced current? Research revealed that an induced current was produced but lasted only so long as the inducing current was undergoing a change.[8] Concluding that the induced current depended not on the mere presence of the inducing current but, rather, on its variation, Faraday went on to examine this point in detail.[9] Not a mathematician, but always an experimenter with a desire for the concrete, he was led to invent a new scheme for representing the magnetic field physically.

The curves traced out by iron filings in a magnetic field suggested to Faraday the concept of a magnetic curve or *line of force*, a curve whose tangent at any point is in the direction of the magnetic field intensity at that point.[10] These lines of force fill all space; every line is a closed curve.[11] If a small nonintersecting curve be taken in space, the lines of force intersecting this curve must form a closed tubular nonintersecting surface termed a *tube of force*;[12] the product of the magnetic intensity and the cross-sectional area is constant along the length of the tube.[13] In virtue of this final property, Faraday conceived all space as being divided up by these tubes, each tube being such that this product was the same for all. Such tubes are termed *unit tubes of force*, and the number per unit area taken normal to them at a given point is a measure of the magnetic intensity at this point.[14]

With this concept as his guide, Faraday pursued his investigations on electromagnetism to the end that he found an electromotive force to be induced in a circuit when (*a*) the current in a neighboring circuit was varied in magnitude, (*b*) a magnet was moved in the vicinity of the circuit, or (*c*) the circuit was moved in the presence of a magnet or an adjacent current. He thereupon concluded that, in every case, induction depended upon the relative motion between the circuit and the magnetic curves in the vicinity, and, further, was able to determine the

[4] Ampère, "Sur la théorie mathèmatique des phénomènes électrodynamiques uniquement déduite de l'expérience," Mémoires de l'Académie des Sciences **6**, 175–388 (1825).

[5] Faraday, "A historical sketch of electromagnetism," Ann. Phil. [2] **2**, 195-200, 274–290 (1821); **3**, 107–121 (1822).

[6] Faraday, "Experimental researches in electricity," Phil. Trans. Roy. Soc. pp. 125–162 (1832).

[7] Faraday, "Experimental researches in electricity" (vol. 1, 1839; vol. 2, 1844; vol. 3, 1855); see pars. 1–5, 26.

[8] Reference 7, pars. 10–11.

[9] Reference 7, pars. 12–26.

[10] Reference 7, par. 114; in modern terminology, *lines of induction*.

[11] Reference 7, par. 3117.

[12] Reference 7, par. 3271.

[13] Reference 7, par. 3073.

[14] Reference 7, par. 3122.

relationship between the direction of cutting of the magnetic curves and the direction of the induced electromotive force.[15] Pressing on, he found in 1832 that currents produced by induction in similar wires of various metals forming circuits of the same geometric shape and position are proportional to the conductivity of the metals.[16] In the light of later knowledge, this may be interpreted to mean that, in essence, induction is the production of a definite electromotive force (emf) which is dependent not on the substance of the conductor but on the interplay between the wire and the lines of force, this emf being produced whether or not the circuit is closed.

Again, he noted that the induced emf is a function of the total number of lines cut in a given period of time,[17] but it was not until 1851 that he wrote "whether the wire moves directly or obliquely across the lines of force in one direction or the other, it sums up the amount of the forces represented by the lines it has crossed"[18] so that "the quantity of electricity thrown into the current is directly as the number of curves intersected."[19] It is to be noted that he did not at any time state definitely—other than for the very special case of a uniform field cut at a uniform ratio[20]—what is now termed *Faraday's law of induction*. We have Maxwell's[21] statement that he himself was uncertain just how clear this was in Faraday's mind in 1832, although he felt that, beyond question, Faraday believed the induced emf to be caused by a change in the number of magnetic lines linked with the circuit.[22]

In addition to this *mutual induction* Faraday[23] was the first to bring out clearly the nature of *self-induction*. William Jenkin[24] and Joseph Henry[25] had noted that, provided the exterior circuit was completed by a long coiled wire, a powerful electric shock could be obtained from a comparatively feeble battery when the circuit was ruptured. Faraday[26] showed that the large momentary current noted on breaking the circuit was produced by a current induced in the circuit and reinforcing the original current as the latter decayed. This is the phenomenon of self-induction; it is described by the same principles as in mutual induction.

Emil Lenz

In 1834 the German physicist Emil Lenz[27] stated what is often called *the second fundamental law of electromagnetism*. This law, correlating the direction of the induced emf in a circuit and the direction of the action of the mechanical force stemming therefrom, was phrased by Maxwell[28] as follows:

> If a constant current flows in the primary circuit A, and if, by the action of A, or, of the secondary circuit B, a current is induced in B, the direction of this induced current will be such that, by electromagnetic action on A, it tends to oppose the relative motion of the circuits.

Franz Neumann

In 1845 Franz Neumann[29] communicated to the Berlin Academy of Science a mathematical theory of the induction of currents. Neumann's[30] own résumé of his first attempts follows in translation:

> If one accepts Lenz's law it follows that where induction is produced by a displacement of a conductor in the presence of an inducing current or magnet, the electrodynamic force exerted by the inducing agent upon the conductor tends always to retard the movement of this latter, and if one notes that the instan-

[15] Reference 7, pars. 114–116.
[16] Reference 7, par. 213.
[17] Reference 7, par. 217.
[18] Reference 7, par. 3083.
[19] Reference 7, par. 3115.
[20] Reference 7, par. 3114.
[21] J. C. Maxwell, *A treatise on electricity and magnetism* (Oxford, ed. 3, 1904), vol. 2, pp. 188–189.
[22] Faraday, reference 7, pars. 217, 238.
[23] Faraday, "On the magneto-electric spark and shock," Phil. Mag. **5**, 349–354 (1834).
[24] In a private letter to Faraday, Jenkin gave the details of his observations and leave to publish them; see reference 7, par. 1049 and pp. 206–207.
[25] Henry, "On the production of currents and sparks of electricity from magnetism," Am. J. Sci. and Arts **22**, 402–408 (1832).
[26] Reference 7, par. 1089.
[27] Lenz, "Über die Bestimmung der Richtung der durch elektrodynamische Vertheilung erregten galvanische Ströme," Ann. der Physik und Chemie [2], **31**, 483–494 (1834).
[28] Reference 21, p. 190.
[29] Neumann, *Die mathematische Gesetze der induzierten elektrischen Ströme* (Berlin Abhandlungen, 1845), pp. 1–17.
[30] Neumann, "Recherches sur la théorie mathématique de l'induction," J. de Mathématique **13**, 113–178 (1848); see pp. 170–171.

taneous induction is proportional to the velocity with which the movement takes place, one arrives at a law expressed by the formula[31]

$$Eds = -\epsilon v C ds.$$

In this formula, ds is an element of the conductor, Eds is the electromotive force developed in the element, v is the velocity of ds, and C is the component, taken in the direction of motion of ds, of the electrodynamic force that would be exerted upon the element if the latter were traversed by a current of unit intensity. The quantity ϵ can be regarded as constant at the instant when induction takes place; considered as a function of time, it is a quantity which decreases with extreme rapidity from the instant that induction takes place.

When the emf Eds is developed in the element ds of the arc s, E is a function not only of s but also of time; however, by supposing that the variations of E with time are incomparably less rapid than the propagation of the electric current in the arc s, one can apply to the induced current Ohm's law, according to which the magnitude of the electric current is equal to the sum of the electromotive forces divided by the resistance of the circuit.

The magnitude of the current induced in the circuit will be, therefore,

$$-\epsilon' \epsilon \int_s v C ds,$$

ϵ' being the conductance of the arc s, and the sign $\int_s$ indicating that the expression $vCds$ must be integrated along the length of arc.

By means of this formula Neumann[32] was able to calculate the induced currents in various special cases. After completing this development, he[33] went on to attack the subject in an entirely different manner:

> I will now show that the current induced in a conductor by a magnet stems only from the change of magnitude that the relative movement of the two bodies brings about in the potential whose Cartesian derivatives represent the three components of the total force exerted upon the magnet by the conductor when traversed by a unit current; thus, one may enunciate the following principle: the change of this potential is the cause of the induction, and if it should be of such magnitude and such manner that this potential undergoes a constant and definite change the induction produced is always the same.

The connection between the afore-mentioned potential and the induction of current can be shown quite easily.[34] Consider two circuits s and s' carrying currents I and I'. The work required to bring these two circuits close to each other from an infinite distance apart is

$$W = II' \oint \oint \mathrm{ds} \cdot \mathrm{ds}'/r,$$

where W is the mutual potential energy of the two circuits and is Neumann's potential function multiplied by II', and where

$$\oint \oint \mathrm{ds} \cdot \mathrm{ds}'/r$$

represents the total number of lines of flux that pass through s owing to unit current in s'. By Faraday's law, the emf induced in s depends only on the variation through s of the lines coming from s'. Thus, Neumann[35] discovered that this potential function supplies all that is needed for an analytic study of the induced electromotive forces in s. Further, it is to be noted that, if unit current exists in both circuits, Neumann's potential function is numerically equal to the *mutual-inductance* of the two circuits, the same following from the definition of this quantity.

Wilhelm Weber

A short time after Neumann made known his theory of induction, the German physicist Wilhelm Weber[36] put forward a theory that unified electrodynamics and electrostatics. This theory was based on certain assumptions originally promulgated by G. T. Fechner,[37] according to which an electric current consisted of identical streams of oppositely charged particles passing through a conductor in opposite directions, like charges being assumed to attract when moving parallel in the same direction, unlike charges attracting when moving parallel in opposite

[31] For a derivation of this formula see reference 30, pp. 114–115.
[32] Reference 30, pp. 113–145.
[33] Reference 30, p. 145.
[34] See, for example, W. R. Smythe, *Static and dynamic electricity* (1939), pp. 304–308; or, F. Neumann, reference 30, pp. 151–155.
[35] Note the previous quotation.
[36] Weber, "Elektrodynamische Maasbestimmungen," Ann. der Phys. und Chemie [2] **73**, 193–240 (1849).
[37] Fechner, "Über die Verknüpfung der Faraday'schen Inductions Erscheinungen mit den Ampère-schen elektrodynamischen Erscheinungen," Ann. der Phys. und Chemie [2], **64** (1845).

directions. With these assumptions Fechner has been able to effect a theory that brought Ampère's and Faraday's laws into conjunction. Weber, adopting Fechner's hypotheses, conceived the total electrodynamic force between two current elements to be the vector sum of the various forces of attraction and repulsion between the charged particles in the two current elements: two electrified particles experienced an electrostatic repulsion or attraction, and, in addition, a mutual force which existed by virtue of the relative velocity between the two particles. Although his formula for the electromagnetic energy associated with two current elements differed from that previously given by Neumann, the two formulas, as Helmholtz[38] later showed, give the same values for the electromagnetic energy associated with *closed* circuits. Previous to this detail, Helmholtz had made important contributions in numerous branches of electric and magnetic theory.

Hermann von Helmholtz

In his classical memoir of 1847 Helmholtz[39] maintained that the potential and kinetic energy of dynamic systems could be converted into chemical, thermal, magnetic, electrostatic and electric forms, these in turn being transformable among themselves. As one illustration of this contention, he considered the action that takes place when a magnet is moved in the vicinity of an electric circuit. Analysis revealed that the energy furnished by the battery equals the sum of the energy dissipated owing to conductor resistance and that communicated to the magnet electromagnetically. Hence, the current is not directly proportional to the emf of the battery, and we have manifested electromagnetic induction. Because Helmholtz, at that time, did not know of the storage of energy in the electromagnetic field, much of the analysis in this paper is wrong;[40] curiously enough, owing to a fortunate, mutually canceling combination of energy exchanges, Helmholtz obtained the right result in the case cited. This problem and the similar one of two electromagnets were later solved rigorously by William Thomson,[41] who largely formulated the true energy theory of magnetic and electromagnetic fields.

William Thomson (Lord Kelvin)

In several preliminary papers Thomson considered various phenomena connected with magnetic fields produced by permanent magnets. In one of these[42] he distinguished between the two vectors **B** and **H**, afterwards termed by Maxwell[43] the *magnetic induction* and the *magnetic force*. In 1851 Thomson[44] had conceived the idea that the energy stored in the magnetic field—this energy he termed the *mechanical value of the current*—due to a current circulating in a closed circuit devoid of emf is equal to the mechanical work done in splitting the circuit into geometrically similar circuits of infinitesimal cross section and bringing these from infinity to form the circuit in question, the individual currents remaining constant throughout this process.

One of Faraday's experiments had shown that a circuit doubled on itself has no stored electromagnetic energy; no spark is manifested when the circuit is ruptured. Starting from this fact, Thomson[45] in 1853 worked out that the energy required to double a circuit in which the current is maintained constant is $\frac{1}{2}LI^2$ and that, therefore, this must be the electromagnetic energy associated with the circuit. He affirmed that L depended only on the geometry of the circuit.[46] Here, for the first time, the concept of inductance

[38] H. von Helmholtz, "Über die Bewegungsgleichungen der Elektricität für ruhende leitende Körper," J. für die reine und angewandte Mathematik **72**, 57–129 (1870).

[39] Helmholtz, *Über die Erhaltung der Kraft* (Berlin, 1848).

[40] Maxwell, reference 21, vol. 2, p. 192.

[41] Thomson, *Papers on electrostatics and magnetism* (London, ed. 2, 1884), pp. 446–447.

[42] Thomson, "A mathematical theory of magnetism," Trans. Roy. Soc. **141**, 243–285 (1851); see par. 48 and 78.

[43] Maxwell, reference 21, pp. 24–25.

[44] Reference 41, pp. 445–446.

[45] W. Thomson, *On the mechanical value of distributions of electricity, magnetism, and galvanism*, Proc. Glasgow Phil. Soc. **3**, 281–285 (1853).

[46] Reference 45, p. 285: "Hence it is concluded that the mechanical value of a current of given strength in a linear conductor of any form is determined by calculating the amount of work against electrodynamic forces required to double it upon itself, while a current of constant strength is maintained in it. The mathematical problem thus presented leads to an expression for the required mechanical value consisting of two factors, of which one is determined according to the form and dimensions of the line of the conductor, irrespectively of its section, and the other is the square of the strength of the current."

appears as a specific entity. In 1860 Thomson[47] defined L as the *electrodynamic capacity* of the circuit, and stated that the energy stored in the electromagnetic field of a circuit, and thence the inductance, could be calculated from the formula

$$(1/8\pi)\int\int\int H^2 dx dy dz.$$

Here, H is the field strength; the integral is extended over all space. Although given for a single circuit in a medium of unit permeability, this formula can be extended to calculate the energy stored in the field produced by any number of circuits carrying current. Hence, in a homogeneous medium of permeability μ, it is possible to calculate the self- and mutual-inductance coefficients of the circuits present.

It is interesting to note that the concept of inductance originally arose from the consideration of the energy stored in the field of a circuit. The definition on the basis of flux linkages follows when it is noted that Thomson's method of obtaining the energy stored in the field of a circuit—and from this the inductance of the circuit—is identical with that used by Neumann to obtain his potential—and hence the mutual inductance of two circuits.

A comparison of the formulas given for the energy stored in the magnetic field of permanent magnets and for the energy stored in the electromagnetic field produced by currents shows that they are identical. Accordingly, the energy stored in the field owing to any combination of permanent and temporary magnets and current-carrying conductors can be calculated. Further, this formula is such that the energy can be considered as distributed throughout the medium with a consequent density at any point equal to $\mu H^2/8\pi$. This concept of distributed energy is a distinctive feature of Maxwell's theory.

James Clerk Maxwell

A contemporary of Thomson was James Clerk Maxwell, first director of the Cavendish Laboratory, who formulated a theory of electromagnetism which, although amplified by his successors, remains fundamentally unchanged to this day. Inspired by Faraday's researches, Maxwell plunged into the task of translating the former's experimental results and ideas into mathematical form. The first major memoir[48] devoted to this undertaking was presented before the Cambridge Philosophical Society in 1855 and was published in its *Transactions* some years later. Maxwell likened Faraday's lines of force to the lines of flow of an incompressible liquid, the induction density corresponding to the velocity of such a fluid. Taking the equation

$$\text{Curl } \mathbf{A} = \mathbf{B},$$

where $\mathbf{A}$ is the vector potential and $\mathbf{B}$ is the induction density, and assuming that Div $\mathbf{A}=0$, Maxwell found that the emf induced in an element of a conductor is proportional to $-d\mathbf{A}/dt$ taken at the element. Further, Maxwell showed that Curl $\mathbf{H}=4\pi\mathbf{I}$; or by Stokes's theorem, the line integral of magnetic force around the curve bounding a surface equals 4π times the current through the surface.

Following this paper, Maxwell proceeded to work out a mechanical conception of the magnetic field. In a series of papers[49] presented in 1861 and 1862, he set forth his vortex theory of the electromagnetic field, the formation of this concept being, beyond doubt, strongly influenced by certain of the vortex researches of Thomson and Helmholtz. In these papers appear most of Maxwell's field equations. The ideas embodied in these papers were unified and extended in the course of the next few years. Finally, in 1864, Maxwell[50] presented to the Royal Society a memoir which marked the completion of the foundations of his electromagnetic theory. He and many others were to amplify and investigate the multitudinous consequences of this theory, but its substance has persisted virtually unchanged.

In the introduction to this paper Maxwell states that, although the "action at a distance" theory developed by Weber and Neumann had

[47] Thomson, "Dynamical relations of magnetism," *Nichols cyclopaedia of physical sciences* (1860), p. 548: "The electrodynamic capacity of a linear conductor of any form is double the mechanical value of a current of unit strength circulating in it."

[48] Maxwell, "On Faraday's lines of force," Trans. Cambridge Phil. Soc. **10**, 26–83 (1864).

[49] Maxwell, "On physical lines of force," Phil. Mag. **21**, 161–175, 281–291; 338–348 (1861); **23**, 12–23, 85–95 (1864).

[50] Maxwell, "A dynamical theory of the electromagnetic field," Phil. Trans. Roy. Soc. **155**, 459–512 (1865).

proved to be exceedingly useful in explaining certain electric, magnetic and electromagnetic phenomena, he yet finds it difficult to accept the assumption of particles acting at a distance with forces dependent on their velocities; he prefers to think of these forces as being produced by actions that go on in the surrounding medium as well as in the bodies themselves, and to explain the actions that take place between separated bodies without having recourse to forces which can act directly over appreciable distances. By considering certain optical, electric and magnetic phenomena, he was led to conceive of an *ether*, of small but actual density, filling all space and permeating all bodies therein, capable of being set in motion and of transmitting this motion from one point to another with finite velocity. From this last property it follows that the motion of one portion of the ether depends on the motion of the remainder, and, the velocity of propagation of the motion being finite, a certain yielding or displacement takes place. Accordingly, the medium may receive and store energy: a kinetic energy, dependent on the motion of its portions and with which the electromagnetic energy is associated; a potential energy, given up when recovering from displacement and with which the electrostatic energy is associated. Propagation of energy through this medium consists of a continual transformation from one form of energy to the other.

From the fact that a body lying in a changing electric or magnetic field or moved through a field experiences a force which can cause charge to flow, can decompose the body or can produce electric polarization, all of which effects tend to disappear when the disturbing force vanishes, Maxwell was led to explain this electromotive force as the agency by which the emf of one part of the ether causes motion in the other. If this emf acts on a closed conducting circuit it produces a current which, in turn, produces heat, the action being irreversible. If it acts on a dielectric it produces a state of electric polarization or electric displacement, this phenomenon being complicated by the conductivity of the dielectric which, although small, may not be inappreciable; and by electric absorption, the phenomena manifest in the residual charge of the Leyden jar.

To summarize: Maxwell's theory hinged on the assumption of an ether, permeating all bodies and modified by their presence; parts of the ether can be set in motion by magnets and electric currents, the motion being propagated out from one part by emf's arising from the elastic connections of these parts, these emf's producing certain displacements or yieldings of the ether, thereby resulting in energy existing as *kinetic energy* due to motion of these parts and *potential energy* due to the elastic displacements. By virtue of this mechanism and provided the principles interrelating the motions of the various parts of the ether are known, the consequences of its motion should be derivable by the laws of dynamics.

Having deduced the nature of his ether, Maxwell went on to consider induction in circuits. If a current exists in a circuit, a magnetic field is produced. Two such circuits will produce a field in which the individual effects are combined: hence, each part of the field is in correspondence with the two circuits, and, therefore, these are in correspondence with each other. One result of this correspondence is the possibility of inducing a voltage in one circuit by varying the current in the other circuit or by moving the first circuit in the field of the second. The laws of induction deduced, Maxwell went on to obtain from them the laws of mechanical action between current-carrying conductors, thus reversing the procedures of Helmholtz and Thomson.

In the deduction of these laws there appear certain coefficients, L_1, L_2, and L_{12}. To these Maxwell gave the name *coefficients of self- and mutual-induction*. He showed that they were functions only of the geometry and relative position of the two circuits[51] and were related to the energy stored in the magnetic field by the equation

$$W=\tfrac{1}{2}L_1I_1^2+\tfrac{1}{2}L_2I_2^2+L_{12}I_1I_2.$$

He indicated how these coefficients could be measured by means of an inductance bridge, gave three methods by which they could be calculated for circuits composed of conductors of infinitesimal cross section and pointed out how, by considering any conductor to be composed of an aggregate of conductors having infinitesimal cross sections, the inductance of a circuit of finite cross section could be determined. Some years

[51] They are, of course, also functions of the permeability; Maxwell was concerned here only with a medium of unit permeability.

later[52] he advanced the concept of *geometric mean distance*, which materially simplified the numerical labor involved in calculating the inductance of circuits composed of parallel cylindrical conductors of finite cross section.

In 1873 Maxwell's[53] comprehensive *Treatise on electricity and magnetism* was published. In this treatise is found the first explicit analytic solution of a problem involving *skin effect;* namely, the calculation of the components of the emf per unit length of an isolated linear cylindrical conductor of radius r which carries an alternating current that distributes itself over the cross section in accordance with the electromagnetic field it produces.[54] In the 1864 memoir[55] Maxwell had shown that the steady-state direct-current inductance of a circuit was not identical with the transient-state inductance and had explicitly stated that the variation was due to the nonuniform current density over the cross section of the cylindrical conductor during the transient period. He now gave an exact solution.

Obtaining for e the expression

$$Ri+L\frac{di}{dt}-\frac{1}{12R}\frac{d^2i}{dt^2}+\frac{1}{48R^2}\frac{d^3i}{dt^3}-\frac{1}{180R^3}\frac{d^4i}{dt^4}+\cdots,$$

where R and L are the direct-current resistance and inductance, respectively, Maxwell[56] pointed out that the terms involving derivatives of order larger than one spring from the fact that the current density is not uniform over the cross section of the conductor. As Rayleigh later remarked, Maxwell evidently did not realize that the alteration of resistance and inductance with frequency might be of practical importance; nor did Maxwell separate the right-hand member of the foregoing equation into components due to resistance and inductance, respectively. This was first done by Rayleigh;[57] however, Heaviside,[58,59] several years earlier, in the course of an exhaustive investigation of the electric and magnetic phenomena associated with a current through long linear solenoids wrapped about cores of different permeabilities, had shown how, as the frequency increased, the current induced in the core tended to concentrate on the outer part of the cross section with a consequent change in the effective impedance of the electric circuit involved, and had given formulas for the circuit impedance as a function of the frequency of the impressed voltage.

Oliver Heaviside and Lord Rayleigh

The series of papers containing the researches just mentioned appeared in *The Electrician* during the years 1884 and 1885. Like much of Heaviside's work, they apparently attracted little attention when first published. Notice was centered upon them, however, when in 1886, in his inaugural address as president of the Society of Telegraph Engineers, Professor D. E. Hughes[60] discussed the results of a series of measurements made upon coils subjected to applied voltages of different frequencies and magnitudes. Hughes interpreted his data as manifestations of phenomena not hitherto known. Heaviside,[61] however, in a rather acrimonious letter, pointed out that some of the experimental data were invalidated by the faulty use of the inductance bridge employed in finding the same, and that the remainder, when properly interpreted, confirmed analytic deductions concerning the change of resistance, inductance and impedance with frequency as published in the papers just mentioned.

[52] Maxwell, "On the geometrical mean distance of two figures on a plane," Trans. Edinburgh Roy. Soc. **26**, 729–733 (1872).

[53] Maxwell, *Treatise on electricity and magnetism* (Oxford, 1873).

[54] Reference 53, pp. 310–312.

[55] Reference 50, p. 509: "But at the commencement of a current and during its variation the current is not uniform throughout the section of the wire, because the inductive action between different portions of the current tends to make the current stronger at one part than the other."

[56] Maxwell, *A treatise on electricity and magnetism* (Oxford, ed. 3, 1904), pp. 321–323.

[57] Rayleigh, "On the self-induction and resistance of straight conductors," Phil. Mag. **21**, 381–394 (1886); see p. 388.

[58] Heaviside, "The induction of currents in cores," Electrician **12** (1884); **13** (1884); **14** (1884–85).

[59] Heaviside, *Electrical papers* (Boston, 1925), vol. 1, pp. 353–416; see especially pp. 368 and 384.

[60] Hughes, "Presidential address," Electrician **16**, 255–258 (1886).

[61] Heaviside, "Note on the self-induction of wires," Electrician **16**, 471–472 (1886). In this letter appears the first specific reference to what has since been called the *proximity effect*, the change in current density over a conductor due to the field of the parallel return. It is of interest to note, however, that Maxwell had previously deduced that the inductance of a given circuit is apparently decreased when current exists in an *adjacent* circuit, while its apparent resistance is increased. See Maxwell, reference 50, p. 475.

These criticisms were shortly accepted as right, although not without some remonstrance on the part of Hughes, whose objections Heaviside[62] was quick to squelch in an even more politely biting letter than the first.

Heaviside was not the only one who criticized Hughes's work. In the discussion following the address,[63] S. P. Thompson, Rayleigh, G. Forbes, W. A. Aryton and others pointed out possible errors in experimental technic and interpretation of data. Subsequently, other criticism was advanced by H. F. Weber[64, 65] and W. Smith.[66]

In the same year Hughes[67] published the results of a second set of researches. His interpretation of them was again challenged by Weber[68] who pointed out divers flaws in Hughes's reasoning.

The publication of Hughes's researches and the ensuing controversy as to their validity and meaning aroused great interest, brought to the attention of engineers the importance of skin effect and proximity effect, especially at the higher frequencies, and centered interest upon the calculation of the effective resistance[69] and inductance[70] of circuits of various shapes. It was but natural, this being before the invention of wireless, that circuits composed of linear parallel conductors were the first to be considered. Thanks to Maxwell, the necessary mathematical theory was available, and the problems could be resolved.

In 1886 Rayleigh published two papers treating such problems. The first of these was concerned with a general dynamical principle regulating the effects of constraints on the motion of a material system. Therein Rayleigh[71] showed that the differential equations characterizing such a system could be interpreted in terms of the parameters characterizing a system of inductively coupled electric circuits, and that the change in resistance and inductance with frequency were natural consequences of the afore-mentioned principle. In the second paper Rayleigh[72] applied the equations of the first paper to obtain a more refined solution of Maxwell's skin-effect problem, giving explicit expressions for the ratio of alternating-current to direct-current resistance and inductance. He also pointed out certain errors in Maxwell's treatise—misprints in the text and certain oversights in analysis.

In the same year appeared the first of a series of papers by Heaviside.[73] the whole comprising an exhaustive investigation of the phenomena associated with the propagation of longitudinal currents in linear cylindrical conductors. These papers contain numerous new formulas for the inductance of various combinations of parallel cylindrical conductors with various ranges of impressed frequency; in addition, the formulas earlier given by Rayleigh are extended.[74] Subsequent to Heaviside's papers appeared many by other authors, treating in detail certain cases which arose with the advance of electrical engineering, and of which a detailed solution was necessary. Of these, one is of particular interest. In his 1889 *Inaugural address* to the Society of Telegraph-Engineers, Lord Kelvin[75] treated,

---

[62] Heaviside, "Note on the self-induction of wires," Electrician **16**, 510, (1886). In this letter Heaviside calls attention to the fact that in January 1885, he had explicitly described how, during the transient state, the current distribution varied over the cross section of a cylindrical conductor carrying a longitudinal direct current; see Heaviside, "On the transmission of energy through wires by the electric current," Electrician **14**, 178–180, (1885).

[63] Hughes, "Inaugural Address: the self-induction of an electric circuit in relation to the nature and form of its conductor," J. Soc. Telegraph-Engineers **15**, 6–25 (1886); discussion, pp. 26–100.

[64] Weber, "Certain critical remarks on Professor Hughes's researches, Electrician **16**, 451 (1886).

[65] Weber, "Critical remarks on the newest discoveries of Mr. Hughes regarding self-induction in metallic conductors," Elec. Rev. **18** ,321–323 (1886).

[66] Smith, "Letter," Electrician **16**, 455–456 (1886).

[67] Hughes, "Researches upon the self-induction of an electric current," Electrician **17**, 71–73, 95–96, 118–119 (1886).

[68] Weber, "Remarks on the second paper of Mr. Hughes regarding self-induction," Elec. Rev. **19**, 30–33 (1886).

[69] This expression was introduced by W. Thomson; see the footnote in his *Collected mathematical and physical papers* (London, 1890), vol. 3, p. 492.

[70] The term *inductance* was coined by Heaviside; see his "Notes on nomenclature," Electrician **16**, 271 (1886): ". . . The inductance of a circuit is what is now called its coefficient of self-induction, or of electromagnetic capacity."

[71] Rayleigh, "Reaction upon the driving-point of a system executing forced harmonic oscillations of various periods," Phil. Mag. **21**, 369–381 (1886).

[72] Rayleigh, "On the self-induction and resistance of straight conductors," Phil. Mag. **21** 381–394 (1886).

[73] Heaviside, "Electromagnetic induction and its propagation," Electrician **17** (1886); **18** (1886–87); **19** (1887); **20** (1887–88).

[74] Heaviside, *Electrical papers* (Boston, 1925), vol. 2, pp. 47, 49, 59, 64, 97, 99.

[75] Thomson, "Ether, electricity, and ponderable matter," J. Soc. Telegraph-Engineers **15**, 6–25 (1886). The mathe-

among other topics, skin effect in cylindrical conductors. In this lecture he introduced the *ber* and *bei* functions and gave a table of computed values for a limited range of arguments.

Several years later Schuster[76] contributed a valuable note on the vector potential. In it he considered the surface conditions that must be satisfied by the vector potential at points in an interface between two mediums of different permeabilities. This topic had not previously been treated in detail.

Conclusion

The men who followed Maxwell and who extended his work—Lorenz, Hertz, Lorentz and others—were concerned chiefly with the electrodynamics of radiation and wave-propagation phenomena. In their work Maxwell's theory appears in various guises, but its substance remains unchanged. Thus, although approximately 80 years have passed since Maxwell's classical memoir was published and although the mechanical interpretation of Maxwell's equations as originally comprehended by Maxwell himself has long since been largely discarded, the mathematical theory contained in that same paper is yet used to obtain the solutions of the complex problems encountered in modern electrophysics and electrical engineering.

While all references cited have been read and while, for the most part, the material presented is drawn from the original sources, I must acknowledge my indebtedness to that valuable work, E. T. Whittaker, *A history of the theory of ether and electricity*, from which I obtained leads to certain of the earlier published references.

matical details of Thomson's solution appear in his *Collected papers in mathematics and physics* (London, 1890), vol. 3, pp. 511–515.

[76] A. Schuster, "Electrical notes: the vector potential," Phil. Mag. **32**, 9–20 (1891).

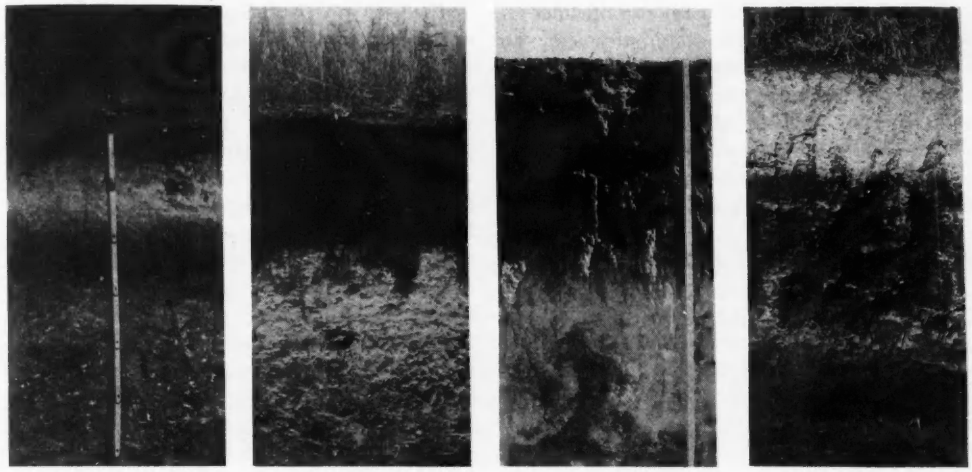
Soil Morphology and Soil Physics

C. C. Nikiforoff
Bureau of Plant Industry, U. S. Department of Agriculture, Washington, D. C.

IDENTIFICATION of the soil types, correlation of the individual soil species and a general classification of the soils are based primarily on certain elemental characteristics which collectively comprise the *soil morphology*. These elemental characteristics include color, texture, consistency, porosity and structure. Some of these, such as texture, are more or less constant, whereas others, such as color, consistency and structure, are subject to considerable changes in different degrees of moisture of the soil.

The taxonomic unit of soil morphology is a *soil horizon*, or a section of the soil body that has acquired individual morphological characteristics different from those of the other parts or sections of the same soil. Therefore, in textbooks dealing with soil morphology, the elemental characteristics are usually referred to the individual horizons. Every soil consists of several horizons which collectively comprise a *soil profile* [(Fig. 1 (*a*), (*b*), (*c*) and (*d*)]. Literally, a soil profile is a two-dimensional homaloidal representation of the vertical cross section of the soil body. In such a cross section the soil horizons appear in the form of bands that extend more or less parallel to the surface of land. These bands, however, represent only the visible edges of the soil horizons which extend throughout the area that is occupied by a given type of soil.

The term *horizon* has a different meaning in soil science and in geology. The horizon of the geologists denotes position; it has no thickness, being merely a stratigraphic level or plane. On the other hand, a *soil horizon* is a layer or bed; it has a certain thickness which is one of its fundamental characteristics no less important than color, texture or consistency of the material of which this horizon is composed. The thickness of soil horizons ranges from a fraction of an inch



(*a*) (*b*) (*c*) (*d*)

FIG. 1. Typical soil profiles: (*a*) Podzol, a forest soil; (*b*) *Chernozem*, a tall grass prairie soil; (*c*) a short-grass prairie soil of the semi-arid regions; (*d*) *Solod*, a meadow soil affected by poor drainage.

to several feet. Only in rare cases is the thickness of any individual soil horizon greater than 3 or 4 ft.

The sequence or arrangement of various horizons in the soil profile, their thickness and the character of the transitions between the adjacent horizons themselves are important morphological characteristics by which various taxonomic units of the soil can be identified. Therefore descriptive soil morphology deals primarily with the arrangement of the soil horizons, their dimensions and such characteristics of their materials as color, texture, consistency, porosity and structure.

A morphology in itself is nothing more than an external manifestation of certain chemical and physical reactions that took place in the original parent material from which the soil horizon has been developed. Behind each morphological characteristic there exists a long story of a certain process. The sum total of morphological characteristics represents what one may call the *soil anatomy*, and the sum total of the processes which lead to the development of these characteristics represents the soil genesis as well as its physiology.

Soil genesis implies a process of gradual transformation of the inert parent material into a definitely organized dynamic soil. A mature soil is a system that is in equilibrium with its environment, the latter including the thermodynamic, hydrodynamic and biodynamic conditions on the surface of the earth. Therefore, soil genesis may be regarded as a process of readjustment of the fresh parent material to the conditions that prevail throughout the medium of pedogenesis.

Moreover, modification of the parent material into the soil is not a continuously progressing process. After reaching a condition of equilibrium with its particular environment the soil remains constant until some changes in environment disturb the equilibrium. The latter may be caused by changes in climate or vegetation, by agricultural practices of men, by erosion of land, and so forth. Morphological characteristics of the soil, however, are neither permanent nor static. They exist as long as the processes upon which they depend continue to operate. As soon as the process ceases to function, the corresponding soil characteristic begins to fade, to degrade and finally vanishes. It follows that morphological characteristics indicate not only the trend of soil genesis but also its present functions. This should demonstrate the significance of soil morphology in general soil science. It is the record of the life-story of the soil. The real value of such a record, however, can be appreciated only after deciphering of the alphabet in which it is written.

To put it another way, the full significance of various morphological characteristics can be understood only if one knows the nature of the processes which lead to their development. One soil is red and another may be brown. Their color is a determining characteristic upon which they may be classified. Yet a recognition of this difference in color is of little value if one does not know why it arises.

A comprehensive investigation of the soil involves at least four fundamental problems. It must answer four questions: what, how much, how and why?

As an independent scientific discipline soil science is rather young. It builds its dominion at the crossroad of geology, biology, chemistry, physics and climatology in which it develops its own contributing departments. Roughly speaking, soil chemistry deals primarily with the problems of *what* and *how much*, whereas soil physics handles the problems of *how* and *why*.

Soil morphology and soil physics are very closely related to each other. Perhaps, the main difference between them is that morphology deals especially with the finished products. It measures, describes, correlates and classifies these products, whereas soil physics is called upon to explore how these products are formed and why they develop. Indeed, no sharp line of demarcation between soil physics and morphology can be drawn because they deal with different aspects of essentially the same objects. In the same way soil physics is intimately related to soil chemistry.

Presumably, every so-called *normal* soil develops from a more or less homogeneous parent material. It may be assumed that such a material originally had a relatively uniform physical, petrographic and chemical composition including a relatively uniform texture, color, consistency and structure throughout the layer which later on was affected by the pedogenic activity.

The differentiation of the genetic soil horizon in which these original characteristics underwent certain changes is due to the uneven intensity of the processes of soil formation in different parts of the parent material and to a different general character of the processes at different depths from the surface of the land.

It seems safe to assume that the decomposition of various minerals and the different means of disposal of various products of such a decomposition are the principal causes of changes in the original material. The products of decomposition may be disposed of in two different ways. They may be carried away by leaching from the place in which they are released, or they may be recombined in the same place into new forms that are more stable against the decomposing agencies than the original minerals from which they were liberated. The materials that are carried away may be either leached from the soil entirely or transported only a short distance and deposited in another soil horizon. The shifting of various constituents from one soil horizon into another may be referred to as their *migration*, and their recombination in place within a given horizon as *segregation*.

Naturally, the fate of the products of decomposition, that is, whether they are able to migrate or become segregated, depends primarily upon their solubility and the movement of the solvent. The common solvent is soil water, usually charged with carbon dioxide. If the solvent moves predominantly downward, as is the case in most of the humid regions, then the translocation of the soluble soil constituents from the upper soil horizons into the lower may take place. If the solvent moves predominantly upward, as in arid regions, then the soluble constituents will be carried toward the surface and will accumulate in the upper soil horizons. In certain instances lateral movement of water—seepage—also may effect the translocation of the dissolved constituents.

The solubility of the products of decomposition depends to a considerable extent upon the reaction of the medium in which they may be liberated or through which they have to pass. Since the reaction usually is not the same in different soil horizons, it may cause a solution of certain constituents in one soil horizon and their precipitation in another horizon.

Obviously, any new morphological characteristic which may be acquired by the soil because of either migration or segregation of one or more of its constituents will be accompanied by changes in chemical and mineralogical composition of the parent material.

The effects of redistribution of various soil constituents due to their migration from one soil horizon into another can be clearly demonstrated by chemical analyses of the whole materials of different soil horizons and by a direct comparison of the data thus obtained with the data on chemical composition of the original material; such a comparison demonstrates either losses or gains of one or another constituent in this or that soil horizon. However, the changes in composition of the original material in various soil horizons which supposedly are demonstrated by such analyses may be only apparent. The true losses or gains of various constituents can be demonstrated only by a calculation of the data obtained by analyses on a common basis. Such a basis may be obtained by assuming for some particular soil constituent a fairly constant value of content throughout the soil profile, which would enable one to recalculate results of analyses upon a *parent material quotient*.[1]

Furthermore, the true losses or gains of various constituents in different soil horizons are not strictly comparable with one another unless they are calculated as *total* losses or gains. The latter can be learned by calculating the true losses or gains upon a *volume factor* of the corresponding soil horizon which represents a relative thickness of such horizons. For example, consider a soil profile in which one horizon is 12 in. thick and another 20 in. thick. Suppose that the first horizon has lost 5 g of sesquioxides per 100 g of original material, whereas the second horizon has gained 4 g of these constituents for the same mass of original soil. Consider a vertical column of this soil of such thickness that it contains 100 g of soil per vertical inch. If such a column is extended throughout the first and second horizons, then the total loss and gain of the sesquioxides from this column will be -60 and $+80$ g in the first and second horizons, respectively. Obviously, the precise value of the volume factor depends upon the thickness of the soil horizons, the density of the soil and the specific gravity of its solid materials. The thickness of the horizons can be determined only by means of a scrupulous observation of the soil morphology, whereas the density and specific gravity of their materials has to be determined by a soil physicist. This demonstrates the importance of soil morphology and soil physics as a basis for calculation and proper interpretation of the data obtained by a soil chemist.

Needless to say, the differences in chemical composition of various soil horizons of a given profile can be attributed to the migration of soil constituents only if this soil has been developed from a thoroughly uniform original material. A great many soils develop from redeposited and more or less stratified parent materials. In such a soil one horizon may develop from a material which originally was different from that of the other horizons. Obviously in these soils the new or *acquired* characteristics, chemical as well as physical, are superposed upon the pre-existing or *inherited* characteristics. Extreme care must be taken in distinguishing between the two types. Here again a soil chemist is practically helpless if his work is not coordinated with the work of a morphologist, a physicist and a petrographer. Unfortunately, many otherwise valuable chemical studies lead to erroneous conclusions simply because the character of the original material and its original dissimilarities in various parts of the soil profile were not taken into account. The investigation of physical properties of the soil coupled with data on mineralogical composition and soil morphology, and only these, can provide the basis for an assumption of one or another character for the parent material.

During the last few decades soil chemists, as well as soil morphologists, placed rather undue emphasis on the role of migration of soil constituents. There has been an almost irresistible tendency to discover in nearly every soil profile the so-called *A*-, *B*- and *C*-horizons, of which the *B*-horizon has been regarded almost universally as an "illuvial." This unwarranted tendency overshadowed the importance of numerous recombinations of the soil constituents without migration to which the common name *segregation* is given in this paper.

The routine total analyses of the whole soil materials cannot demonstrate the effects of

[1] As used in this paper, the expression *parent material quotient* indicates a volume or weight equivalent of a certain amount of the original material. For example, let us take 100 g of the original material. This amount of material may loose 15 g in one soil horizon and gain 8 g in another horizon. Therefore, the parent material quotients in these horizons will be 85 and 108 g, respectively.

segregation; the average chemical composition of a fundamentally reorganized soil material may be essentially the same as that of the original material. The effects of segregation can be demonstrated only by parallel analyses of different parts of fractions of the materials from various soil horizons. For example, these materials can be divided into colloidal and noncolloidal parts. Again, the soil can be divided into sand, silt and clay or into any other mechanical fractions. A given mechanical fraction may be divided into parts composed of heavy and light minerals, and so forth. Each of these parts or fractions must be analyzed individually.

Here is a simple example. Suppose that the whole material of a given soil horizon and the whole original material have exactly the same percentage of silica, for example, 80 percent. But parallel fractional analyses indicate that 80 percent of the total silica in the parent material is held in noncolloidal parts of the soil, and 20 percent in colloid, whereas, in some particular horizon, only 60 percent of the total silica is in the noncolloidal part of the horizon and 40 percent in colloid. This indicates the transfer of 20 percent of total silica from the noncolloidal part of the soil to the soil colloid, that is, segregation of silica by the colloid from the noncolloidal part of the soil. Such results can be obtained by the calculation of the data of fractional chemical analyses on the basis provided by physical or mechanical analyses.

Again, the direct data obtained by these analyses demonstrate only apparent effects of segregation. The true effects can be learned only by recalculation of these data upon the parent material quotient. Such a recalculation would not be so important, perhaps, if the formation of a new soil horizon had been due entirely to the recombinations of its constituents in place; for in this case, the parent material quotient would be affected only by changes of the bulk density and specific gravity of the original material. There is hardly any soil, however, whose development is due exclusively to either migration or segregation of various constituents. Therefore, a comprehensive chemical investigation of the soil must consider both effects and evaluate each of them separately.

So far we have dealt predominantly with the chemical part of soil investigation. An attempt has been made to present in a most general way what is going on in the parent material during its modification into the soil and the means by which the questions "what" and "how much" can be answered. This subject has been treated at some length in order to present a clearer picture of the objectives with which the problems "how" and "why" are concerned. The special work of the soil chemist probably ends with the last calculation of the data regarding "what" and "how much," and the special work of the soil physicist begins at the very same point.

It has been pointed out that the essence of the transformation of the parent material into the soil is the establishment of a condition of equilibrium between this material and the environment which prevails on the surface of land. It is assumed that the fresh parent material is not in equilibrium with this environment. Instead, it is in equilibrium with the environment which prevails at some depth from the land surface and which is characterized by relatively constant moisture and temperature and is affected very little, if at all, by living organisms. The equilibrium between mineral material and such an environment is relatively static. The changes of this equilibrium are due only to certain slow geological processes of general weathering, and even these are absent at a great enough depth.

Living organisms and variability of moisture and temperature are the outstanding characteristics that differentiate the environment on the land surface from that below a certain depth. This difference of conditions in the two mediums and consequently different states of equilibriums in each system give the general answer as to why the uppermost part of the parent material cannot rest unchanged but undergoes reorganization in the soil.

Each system consists of two components: a relatively passive mineral framework of the soil and a combination of dynamic factors—including living organisms, heat and moisture—which act upon the first component.

Insofar as the formation of any particular soil is concerned, the dynamic part of the system remains in a state of relatively constant kinetic equilibrium. Any deviation from such a state would detract the process of development of the

soil toward a somewhat different end-product. Therefore, an average biological pressure,[2] average temperature and average moisture must be taken as relatively constant throughout the period of development of any particular soil. These averages, naturally, are the results of interpolation of different values in each system. At the depth at which the parent material remains unchanged, biological pressure is practically zero and moisture and temperature are nearly constant, whereas, near the surface, all these factors are subject to more or less rhythmic diurnal and annual fluctuations.

It is a passive solid component, especially the mineral framework of the soil, that is subject to readjustment in the systems progressing toward a kinetic equilibrium in the soil and a potential equilibrium in the parent material. There is no sharp line of demarcation between the zones occupied by these two systems and hence no abrupt boundary between the soil and its unmodified parent material. The biological pressure, the average temperature, the average moisture and, especially, the magnitudes of diurnal and annual deviations from the means are not the same throughout the zone of pedogenesis. Because of this the original material does not change evenly throughout the soil body and the soil acquires a profile composed of several different horizons. Each horizon develops in a specific environment somewhat different from the environments in which the other horizons are formed and, consequently, each horizon attains a specific equilibrium which determines its individual morphological face.

In this paper we shall not be concerned with living organisms, although they represent the most influential factor of soil formation. This particular branch of soil science relates more closely to biochemistry than to physics. Soil physics deals especially with heat and water relationships in the soil.

Variations of temperature and moisture in the soil are usually spoken of as *soil climate*. Figuratively speaking, soil climate is an echo of the climate of the atmosphere. The climatic elements, of which temperature and precipitation are the most important, do not cease abruptly as they strike the surface of the land, but penetrate for a short distance below the surface. They continue to fluctuate diurnally and annually below the land surface in much the same way as in the air, although with a rapidly decreasing speed of transfer of heat and water from one level to another and with a rapidly shrinking range of deviations from the means.

The daily maximums of the temperature in the air send into the soil the waves of relatively high temperature separated from each other by the waves of somewhat lower temperature which originate from the daily minimums in the atmosphere. The decrease in the speed with which these waves penetrate the soil leads to the retardation of the daily and annual maximums and minimums at different levels of the soil. At a certain depth from the surface the diurnal maximum may occur simultaneously with the diurnal minimum at the upper layer of the soil. To put it another way, the daily maximum in the former layer of the soil may occur several hours later than in the latter. At some greater depth the annual maximum may arrive in January or even later, whereas near the surface of the soil it is more likely to occur in July or August.

Since the transfer of heat is always from the point of higher to the point of lower temperature, a layer of the soil that has acquired its diurnal maximum temperature must send a part of the accumulated heat back upward as soon as the layer above it is chilled to a lower temperature. Thus the quantity of heat that flows downward gradually decreases and the amplitude between the diurnal extremes shrinks to a vanishing point at the base of the soil. The transmission of heat also, of course, depends upon the thermal conductivity and thermal capacity of the medium.

Soil climate refers not exclusively to the rhythmic fluctuation of the temperature and humidity of the air in the soil but to the fluctuation of temperature and moisture in the whole system which is composed of three phases—solid, liquid and gaseous. The solid part of soil is an unconsolidated heterogeneous mixture of organic material and minerals, the individual particles of which range in size from ultramicroscopic clay

[2] The grades of *biological pressure*, as used throughout this paper, refer to the relative amount of active living matter whether in weight or volume per unit area, for example, in tons or cubic feet per acre.

minerals to large fragments of rocks. The liquid part is water carrying variable amounts of dissolved material, including salts, carbon dioxide, organic matter, and so forth. The gaseous part represents the soil air with admixtures of different gases, such as carbon dioxide and ammonia.

The proportions of the solid and nonsolid matter in the bulk of soil vary according to its porosity. The proportions of soil water and soil air in soil of any given porosity range from a nearly complete saturation of the pore space with water to a nearly complete desiccation of the soil, although, under natural conditions, probably no soil can be deprived of all its air or water.

The mineralogical composition of the solid framework of the soil varies from an almost pure assemblage of one mineral, such as quartz, to an extremely complex mixture of many different minerals of different specific heats and thermal conductivities. An assemblage of the same minerals may differ in packing at different places and, consequently, in porosity and in pattern of the pore space. Furthermore, the same assemblage of solid particles may acquire a different structural arrangement in the soil and this may vary with changes of moisture.

This variety of conditions affecting either directly or indirectly the flow of heat through the soil and the quantity of heat which may reach this or that horizon leads to a great variety of soil climates. Therefore a relatively monotonous and uniform blanket of the atmospheric climate becomes a multicolored patchwork as soon as the elements enter the medium of the soil. This variety of the thermal and hydrologic régimes leads to the development of different soils in areas characterized by a relatively uniform atmospheric climate. The differences in soil climate may be small and inconspicuous at any given moment, but since they persist during long periods of time consumed by adjustments of the parent material to the environment, their effects on soil morphology may become rather spectacular.

Exact data on the diurnal and annual flow of heat through the natural soils are very meager. The few and rather casual observations which have been reported are rather amateurish as regards both methods and equipment. Investigations of the thermal conductivity of various soils materials are more numerous and methodologically more satisfactory. Almost all these investigations, however, were dealing, not with natural soils, but with samples of the soil materials that had been dried, crushed, sifted and repacked. It is not yet precisely known how much such operations distort the character of heat transmission in the soil, although it may be expected that the distortion is drastic. Since, in soil development, even a slight deviation from the normal condition may lead to a remarkably strong effect, the data obtained by these studies of artificial soils may have only a relative value. The thermodynamics of the *natural* soils and of their formation are still the "terra incognita" which awaits its explorer.

Water is another outstanding element of the soil climate. There is a certain, although remote, similarity between the movements of water and heat through the soil. Indeed, changes in soil moisture lack the rhythmic regularity characteristic of the flow of heat because of a rather sporadic supply of water.

Movement of water in the soil is particularly important because water is the principal carrier of the mobile soil constituents. It is doubtful whether any appreciable migration of such constituents would take place in the permanently dry soil. Water carries various materials either in suspension or in solution. It drops its cargo either because of the oversaturation of the solution or because of reactions between the dissolved materials which may lead to the precipitation of the new insoluble compounds or to their coagulation. The soil itself may retain certain dissolved matter by absorption and some suspended material by acting as a filter.

Movement of water through the soil is possible because the soil is porous. The volume of pore space ranges in different soils from about 30 to 60 percent of the bulk. In some soils a relatively low porosity may be represented by comparatively large interconnected voids, whereas in the others a high porosity may be represented by a dense network of very small voids. It is customary in soil science to speak about a "capillary" and "noncapillary" soil porosity, although such a nomenclature is open to criticism.

In every soil a part of the porosity is represented by the capillary voids and another part

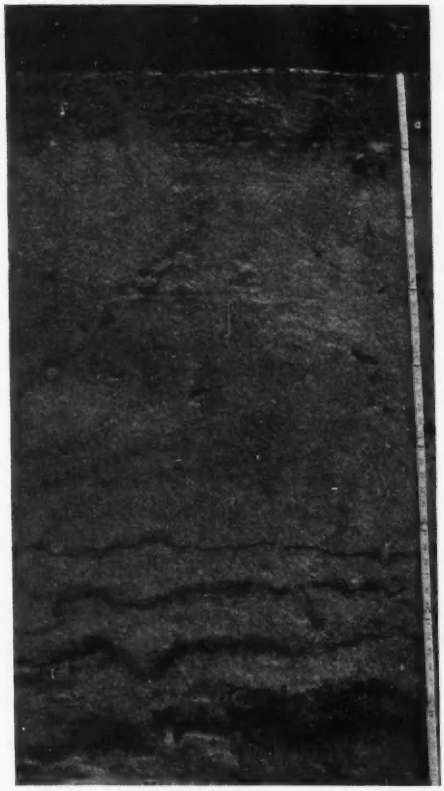


FIG. 2. Precipitation of iron oxides and formation of the thin layers of ortstein (iron cemented sand) in the lower part of the soil profile.

by the noncapillary. The ratio between these parts ranges widely and may vary in a given soil with the moisture content. Many soils shrink and crack on drying, thus opening a network of large voids and increasing the compactness of the crumbs of soil separated from one another by the cracks.

The pattern of the pore space and the range of its variability under different conditions are perhaps more important as regards the receiving, holding and transmission of water than the mere volume of the pore space. It should be noted that artificial treatment of the soil samples destroys the natural pattern of soil porosity and changes the condition of a water transfer even if packing would bring the sample to a density similar to that of a natural soil.

Water moves through the soil because of its weight, capillary action, hydrostatic pressure and, in the vapor state, convection. To these may be added transpiration of soil water by plants and evaporation from the surface of leaves, but since we are omitting consideration of organisms, these causes of water movement will not be discussed in this paper.

Most of the water entering the soil from the surface comes, of course, from rain and snowfall. Some comes from water vapor condensed from the air. A much smaller, perhaps negligible, amount of water comes from below the soil surface. Only certain *hydromorphic intrazonal* soils which develop in poorly drained areas are affected by abundant ground water supplied by seepage or hydrostatic pressure. Except for these hydromorphic soils, the capillary effect and evaporation are the principal causes of rising soil moisture, whereas the force of gravity is the principal cause of the downward movement of soil water. On the basis of the effects produced by gravitational and capillary forces, four different states of soil water are recognized—gravitational, capillary, hygroscopic and combined water. To these, water vapor may be added.

The speed with which water flows through a particular soil depends upon the permeability of this soil to water and its capacity to hold its water against one or another force. Both these conditions are determined by the total surface area of the solid particles and by the soil porosity, especially the size and shape of the voids. Naturally, gravitational water moves faster through relatively large voids, whereas the capillary action takes place in the small channels. This again demonstrates the importance of investigation of the behavior of water in a natural soil rather than in artificially packed samples of the soil material.

This paper is not concerned with the fundamental laws of the soil and water relationship. We are interested now primarily in the effects of movements of soil water on soil formation and, especially, on soil morphology.

Leaching, or a transfer from the upper to lower level, represents the most common form of migration of the mobile soil constituents. Various salts, free oxides and in some instances,

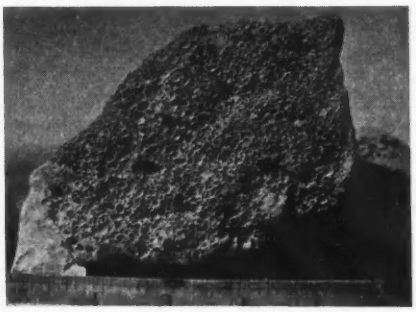


FIG. 3. Vesicular porosity of the upper horizon of the semi-arid soil.

perhaps, colloid as a whole may be shifted from one level to another by leaching. Obviously, the transfer depends chiefly upon the movement of gravitational water in relatively large voids.

Two different forms of leaching may be recognized. One of them is an exceedingly slow geological leaching which removes the soluble products of rock weathering and probably is responsible for their accumulation in the ocean. Geological leaching affects the entire zone of weathering. It accompanies the formation of the unconsolidated *parent material* from which the soil may be developed. It is a continuous process and the soil, being a part of the crust affected by weathering, also is subject to this leaching.

Another form of leaching consists of a downward migration of various soil constituents and their deposition within the soil profile. Most of the rain water which enters the soil fails to ooze through the soil very far beneath the surface. The water front advances downward slowly and with rapidly decreasing speed. Since the influx of water is not continuous, sooner or later the wave of maximum moisture breaks away from the surface. The intervals between the periods of intake of water are usually characterized by rising temperature and evaporation which stimulate a capillary pull toward the level of vaporization. Vaporization creates an upper water front which follows the lower one owing to a gradual desiccation of the soil. Sooner or later the capillary pull exceeds the force of gravity and reverses the direction of movement of water from the soil.

Evaporation, naturally, leads to the concentration of soil solutions and to precipitation of the solutes. However, it does not bring the dissolved matter back to the points from which it originated but drops it in different places. The deposition probably takes place more frequently on the level at which the lower water front has ceased to advance. The examples of such a relocation are many. One of them is evident in Fig. 2, which shows a series of irregular dark streaks near the base of the soil profile. These streaks represent the cuts of irregular crustlets formed by precipitation of iron oxide and some other compounds which have cemented the loose sandy soil into the thin brightly colored crusts. Presumably, they mark the position of the lower water front in different years. Once formed, these clogged sheets begin to act as filters for the following waves of descending solutions and to grow in thickness. In some places they may merge into a continuous soil horizon.

Since gravitational water moves faster through the large voids, it naturally fills them first and then spreads into the smaller pores. These large voids form a network of cracks which break the soil into lumps of various shapes and sizes.

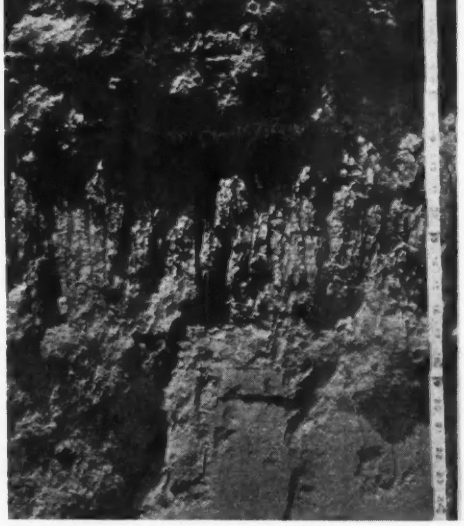


FIG. 4. Prismatic structure of the *B*-horizon in arid short-grass prairie soil.

FIG. 5. Shrinking and cracking of the soil on drying open a net of large voids.

Although some air is trapped inside the lump when the network of voids fills with water, most of the air tends to escape from the surface. In some soils covered by a scant vegetation the tramping of the surface by drops of rain produces a compact crust which obstructs the escape of air bubbles from the soil. A rise of temperature following the rain causes the air bubbles to expand, which leads to the development of a peculiar *vesicular* form of soil porosity, an example of which is shown in Fig. 3.

The intake of water may be accompanied by a considerable swelling of the soil, whereas the desiccation of such a soil causes it to shrink and crack. In some horizons the soil material breaks down into roughly prismatic clods such as those which are shown in Fig. 4; in other horizons it may split horizontally into laminae of various thickness and in still others it breaks into irregular blocks of various sizes (Fig. 5). Each of these types of fragmentation is characterized by a corresponding pattern of the coarse voids which affects differently in each case the flow and transmission of heat and water through the corresponding soil horizons.

Some spectacular effects on soil morphology are produced by the freezing of water in the soil. An example of this is the soil shown in Fig. 6, which froze while in a thoroughly wet condition. The black horizontal streaks represent the thin lenses of pure ice, between which are compressed layers of the soil.

Hydrolysis is one of the most important reactions which alter the parent material. The effectiveness of this process very likely depends to a considerable extent on the duration of a direct contact between the soil and water. The uppermost soil horizon is wetted more frequently and, perhaps, is traversed by a greater volume of water than the lower horizons (except for soils that are affected by seepage). However, it gives up its water more quickly and usually is desiccated more thoroughly than the lower horizons. Therefore, the time during which soil and water are in contact may be shorter in this horizon than in the others.

Obviously, the length of time during which the soil may hold its water depends upon the rates of percolation, evaporation and transpiration by plants. The difference in duration of the soil and water contact in various soil horizons may be either great or small during a particular year, but in the long run it is sufficient for the development of strikingly dissimilar soil horizons from an essentially uniform original material.

These are but a few, although perhaps the most conspicuous, examples of morphological characteristics of the soil and of the physical factors of their development. Many other less spectacular, but by no means less important, examples could be added to this list.

Experimental agriculture demonstrates that some of these characteristics are beneficial as regards the principal predestination of soil in nature—its function as a sort of a clearinghouse between the organic and inorganic worlds—whereas the other characteristics are rather negative, at least from the practical viewpoint of man.

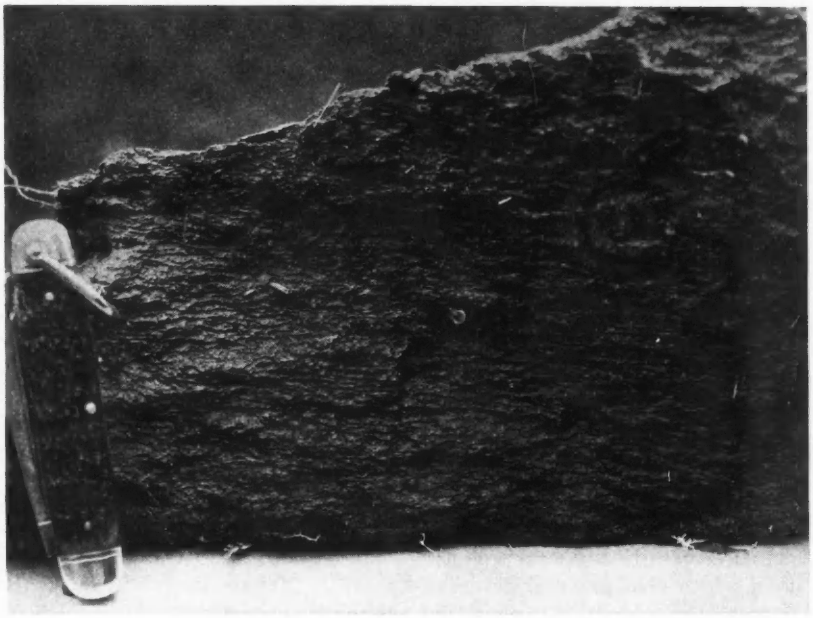


FIG. 6. Effect of freezing of the wet soil. The black streaks represent the thin lenses of ice which split the soil horizontally. (About two thirds natural size.) [Photograph by courtesy of Grant A. Mickelson.]

Man's interest in soil is not mainly academic; he looks upon the soil as a principal source of his food and his wealth rather than with a scientific curiosity. Experience in agriculture teaches him that, although a potential capability of the soil to yield crops and to support the living world is great, indeed, the vital functions of this soil can be ruined by improper management. No other natural resource is subject to such a long and equally drastic treatment and exploitation. Thus through bitter experience he learns the necessity of careful preservation and skilful development of the useful characteristics of the soil and of a cautious control of the others. Needless to say, none of his efforts can be very successful unless he knows how these characteristics have been developed in the first place, what they indicate and what their dynamics are.

Soil physics is not much younger than general soil science. The number of valuable contributions to this discipline is already large, and the flow of new studies continues to increase. Soil physics holds a prominent place in the curriculums of agricultural schools and colleges; textbooks have been written and published. It should be noted, however, that the so-called soil physics of today still represents to a considerable extent a somewhat amateurish combination of descriptive soil morphology and rather unsystematic experimental data regarding various physical properties of the *soil materials*. These materials in most instances are represented by dead soil samples, crushed, sifted, and repacked according to the purpose of the experiment. Obviously, such samples have little in common with a natural soil from which they were taken.

Soil physics should operate with undisturbed natural soil and its physical properties and dynamics. This is the only gate through which one may hope to approach the solution of the problems of how the soil has been formed, why it has acquired one or another morphological face and what is the significance of its morphological characteristics.

Some Interesting Aspects of the Impact Ball Apparatus

SEVILLE CHAPMAN
*University of Kansas, Lawrence, Kansas**

A POPULAR demonstration piece used to illustrate Newton's third law of motion consists of a row of similar steel balls which just touch one another and are suspended by bifilar suspensions so that all the balls swing in the same plane (Fig. 1). Sometimes one or more of the end balls is of larger mass than the rest of the balls in the row.

The motions of the balls when all of them have the same mass m are easily explained by assuming perfect elasticity and employing the principles of conservation of momentum and energy stated in the forms

$$\sum mv = \sum mu, \tag{1}$$

$$\sum \tfrac{1}{2}mv^2 = \sum \tfrac{1}{2}mu^2, \tag{2}$$

where u and v are the velocities of the balls before and after impact, respectively. When the masses of two balls are the same, Eqs. (1) and (2) show that their velocities must be exchanged on impact. In a long row, if four balls are drawn out and then released, four balls fly out from the other end after the impact, all the other balls remaining at rest.

Suppose now, the four balls that are drawn aside originally are replaced by a single ball of mass exactly four times that of the smaller balls. Equations (1) and (2) would still be satisfied if four small balls would again fly out after impact with a velocity equal to that of the large ball before impact. If one tries the experiment, it very definitely does not work that way. Why not?

One's first impulse is to say that the balls are not perfectly elastic, and thereby their failure to follow the expected course is explained. But if one measures the coefficient of restitution e of the balls (the ratio of the relative velocity of two balls after impact to their relative velocity before impact) one finds it to be about 0.99 (Fig. 2). This is too small a discrepancy from perfect elasticity to explain the "jumble" of balls that one finds experimentally.

A second effect to be explained is the following one. Suppose that only three balls are used (for convenience), and that one of the end balls has a mass several times that of the other two. It is found that if the two small balls are drawn aside and released so as to collide with the stationary large ball, a fascinating cyclic motion results. After every two (or four) multiple impacts, the two little balls will fly out together with their original amplitude. The point to be explained is that sometimes there are two collisions to the cycle and sometimes four. It is important to notice that in either case the original conditions of energy and momentum are identical, and yet the same motion does not always occur.

Both of these effects can be explained on the basis of sequence of impacts. If Eqs. (1) and (2) are used to solve for v_2, the velocity of one of the balls after impact, in terms of m_1 and m_2, the masses of the balls, and u_1 and u_2, their velocities before impact, one obtains

$$v_2 = [2m_1u_1 + (m_2 - m_1)u_2]/(m_1 + m_2). \tag{3}$$

Now if one makes a *step-by-step* calculation,

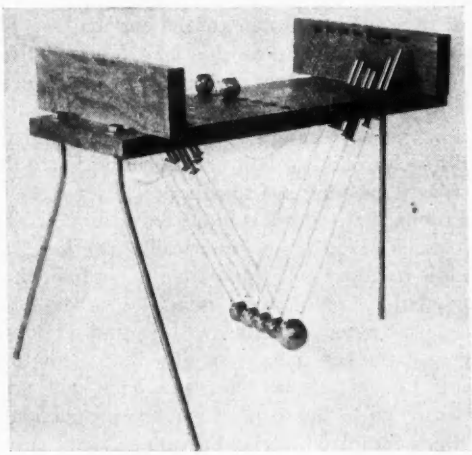


FIG. 1. Impact ball apparatus. Note the adjustable stops by which the effective lengths of the pendulums can be controlled.

* Now at Stanford University, Stanford University, California.

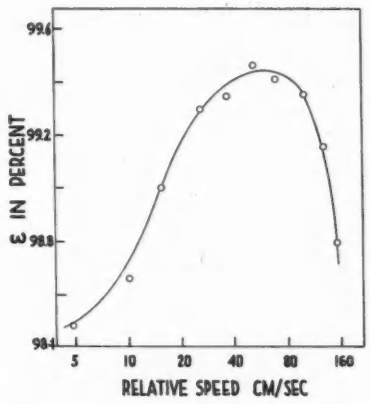


FIG. 2. Coefficient of restitution of steel balls as a function of their relative speed before impact. The values have been corrected for air friction. It is imperative that the balls be extremely clean and well polished if such high values of e are to be obtained.

using Eq. (3), going down the whole line of balls in the first instance and figuring each impact successively, he will arrive finally at a situation where all the balls are separating. From the measured amplitudes of swing, one can check the calculated velocities, and will find that the agreement is quite good.

The second effect with three balls is the simplest case in which there can be any differences in the order of impact.[1] For instance, in a given multiple collision of the three balls, the middle ball may collide first with the end ball on the left, or with the one on the right. The resulting motion will depend on which impact occurs first.

Let us suppose that we allow the big ball of mass 4 units to hang at rest, and that two little balls each of mass 1 unit are drawn aside (to the left, say) and released together. The sequence of collisions that follows is indicated in Fig. 3. The two balls swing down practically together, the middle ball first colliding with the big ball, then immediately rebounding into the other little ball, exchanging velocities with it, and rebounding into the big ball. A *step-by-step* calculation using Eq. (3) shows that the balls will now separate with amplitudes roughly proportional to those indicated in the diagram.

[1] The case of two balls of unequal masses has been discussed by Lemon, Am. J. Phys. (Am. Phys. T.) **3**, 36 (1935).

On the next multiple collision, *if* the first impact is between the middle ball and the *big* ball, the sequence of impacts will be the same as in the first collision, but in reverse order. A step-by-step calculation using Eq. (3) shows that after three impacts the two little balls will fly out together with their original speed so that their amplitude of swing will be the same as their original amplitude.

The process then repeats, and we have a "two-collision" cycle. We shall represent the impacts of the middle ball with the big ball and with the little ball by the symbols B and L, respectively, as in Fig. 3.

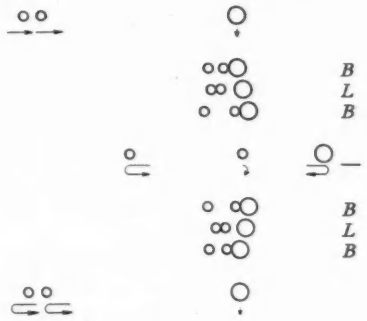


FIG. 3. The complete two-collision cycle.

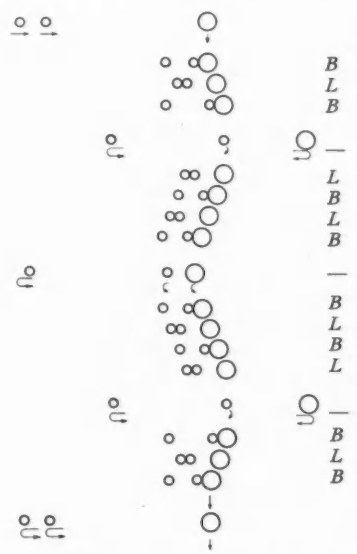


FIG. 4. The complete four-collision cycle.

Now suppose that, on the second collision, the middle ball collides first with the little ball. In that case, a step-by step calculation [Eq. (3)] shows that the events indicated in Fig. 4 will occur. As before, the first collision is a *B-L-B* one, but now the second one is a four-impact collision, *L-B-L-B*. When the balls separate after the first collision, the little end ball will have an amplitude about 30 percent greater than its original amplitude.

If, on the next multiple collision, the first impact is between the middle ball and the big ball, the events of the first two collisions will be carried out, but in the reverse order, so that after four multiple collisions two little balls will fly out to their original amplitude. The process will repeat and we have a "four-collision" cycle. Students always are fascinated by these cycles.

To put these conclusions to experimental test, the balls were connected by flexible wires through an electric circuit to a cathode-ray oscillograph in such a way that when the middle ball collided with the big ball it completed a circuit which produced a slightly larger vertical deflection of the oscillograph spot than was produced when the middle ball collided with the little ball.[2] By setting the sweep circuit of the oscillograph so as to draw the spot rapidly across the screen horizontally, the time during which the multiple collision lasted (about 1/50 sec) was spread out over about $\frac{1}{2}$ in. of screen.

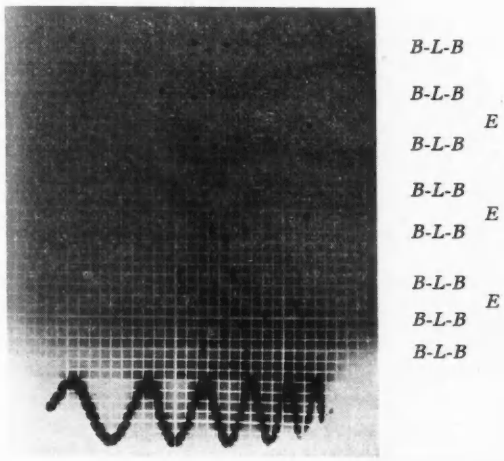


FIG. 5. Photograph of the oscillograph screen showing four complete two-collision cycles.

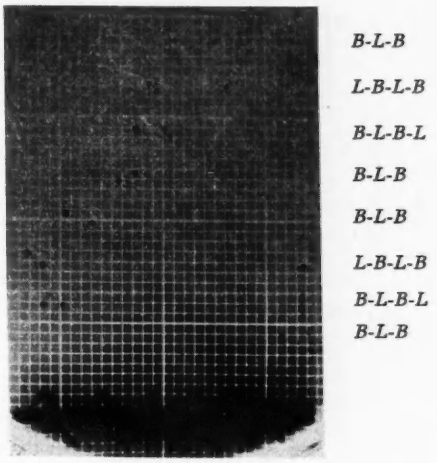


FIG. 6. Photograph of the oscillograph screen showing two complete four-collision cycles.

By noting the order in which the "bigger or littler" deflections were produced, it was possible to note the order in which the middle ball collided with the big ball or with the little ball. In photographing the oscillograph screen an additional electric circuit was included which gradually reduced the amplitude of the vertical deflections, so that as many as eight or ten multiple impacts could be recorded on a single film. In this way the complete sequence of events in several two- or four-collision cycles could be obtained on the one film. Finally a timing wave from an oscillator of adjustable frequency was put on the film for reference.

Figure 5 is a photograph of the oscillograph screen obtained for the two-collision cycle. Again the symbols *B* and *L* stand, respectively, for the impacts of the middle ball with the big ball and with the little end ball. The symbol *E* stands for the extraneous impact sometimes observed when the two little balls fly out to their original amplitude at the end of a cycle, and just happen to touch.

Figure 6 is a photograph of the oscillograph screen obtained for the four-collision cycle. In this case the synchronization between the

[2] The voltages used were such as to preclude sparking; see Loeb, *Electrical discharges in gases* (Wiley), pp. 413 ff.

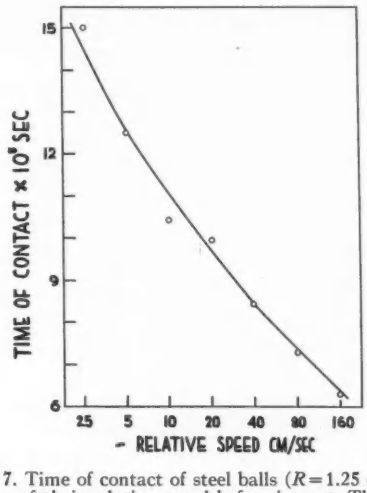


FIG. 7. Time of contact of steel balls ($R=1.25$ cm) as a function of their relative speed before impact. The points are experimental, but the solid line is a theoretical inverse fifth power curve. Since the exact values of the elastic constants of the balls are unknown, the theoretical curve is adjusted to fit one of the experimental points. The constant k of the inverse fifth power law for this adjusted curve has a value 7 percent lower than that calculated from tabulated values of the physical constants.

pendulums and the oscillograph was not quite exact, but it is seen that the cycle is *B-L-B*, *L-B-L-B*, *B-L-B-L*, *B-L-B* as predicted.

Verification has also been obtained for an eight-collision cycle. In this case the events are *B-L-B*, *L-B-L-B*, *L-B-L*, *B-L-B-L*, *L-B-L-B*, *L-B-L*, *B-L-B-L*, *B-L-B*.

While the order of impact can be partially controlled by varying the amplitude of swing, it is most easily controlled by varying the lengths of the pendulums. This can be done without changing the positions of the balls by moving the adjustable stops shown in Fig. 1.

For the two-collision cycle the pendulum of the big ball should be somewhat shorter than for the other two (perhaps $\frac{1}{2}$ in. shorter—the value is not critical). For the four-collision cycle, the pendulum of the big ball should be somewhat longer than for the other two. It takes considerable patience to adjust for an eight-collision cycle. The pendulum of the little end ball should be shorter than that of the middle ball, and the pendulum of the big end ball should be longer. The differences in length are only a few tenths of an inch and are *very* critical.

When only two balls are used, it is interesting to inquire how long they are actually in contact during impact. Hertz[3] predicted that for two equal balls of radius R approaching each other with a relative velocity u, the time of impact t should be $kR/u^{1/5}$, where k is a constant for any particular material, being a function of the density, Young's modulus, and Poisson ratio of the material.

With the oscillograph it was a simple matter to measure the time of contact by increasing the frequency of the sweep circuit so that the writing speed of the spot was about 5000 in./sec. The results are shown in Fig. 7 for the only case investigated, that of two steel balls of radius 1.25 cm. It is seen that the inverse fifth power law is quantitatively verified. The surprising thing is that the actual times of contact are so short, being about 10^{-4} sec. This time is about 20 times as long, however, as the time of transmission of a pressure wave through a ball (comparable to the time for the transmission of sound through it) so that the time of contact is long compared with the period of elastic vibration of a ball. No evidence of possible elastic vibration was obtained from the oscillograph.

The author wishes to thank Mr. William Bush for his assistance in making some of the measurements.

[3] Hertz, "On the contact of elastic solids" (1881), from *Miscellaneous papers* (Macmillan, English tr., 1896), pp. 146–162; Poynting and Thomson, A *textbook of physics, properties of matter* (Charles Griffin, ed. 3, 1905), chap. X.

SCIENCE alone can give us a true conception of ourselves and of our relation to the mysteries of existence. Only the sincere man of science—and by this title we do not mean the mere calculator of disturbances, or analyzer of compounds, or labeler of specimens; but him who through lower truths seeks higher and eventually the highest—only the genuine man of science, we say, can truly know how utterly beyond not only human knowledge but human conception is the universal power of which nature, and life and thought are manifestations.—HERBERT SPENCER

The 1941 Summer Engineering Defense Training Program of the Pennsylvania State College

MARSH W. WHITE
Pennsylvania State College, State College, Pennsylvania

THE Pennsylvania State College during the summer of 1941 gave, as part of the Engineering Defense Training (EDT) program sponsored by the United States Office of Education, a course in introductory engineering subjects for qualified high school graduates who were not planning to enter college in the fall. The course was administered by the Extension Services of the College. The initial enrolment was about 3000 students, at 95 centers throughout Pennsylvania.

PROGRAM AND CURRICULUM

The objective of this program was to provide high school graduates with introductory engineering training so that they would be better prepared to aid in American defense production. No college credit was granted for the work, although a certificate was given to those students who completed the courses satisfactorily. The program was designed to assist in fitting these students for:

(1) Positions in industry and in many departments of the Federal Government which are urgently in need of large numbers of inspectors, designers, draftsmen and other technologists with semi-technical training;

(2) Further advanced training in other EDT programs; the summer program qualified the student to continue his study in more specialized engineering, science and management courses, such as those now being offered in ESMDT evening programs;[1]

(3) More rapid advancement after employment; for those going into industry following the summer work there was the probability of more useful service, rapid advancement and greater responsibility.

All students took the same work, in subjects treated from an engineering standpoint. The course required 40 hrs per week in scheduled activities. During the first five weeks the student's program was as follows:

Chemistry, 5 hr/wk, 2 recitations and 1 3-hr laboratory period;
Drafting, 5 hr/wk, 1 3-hr period and 1 2-hr period;
Physics, 8 hr/wk, 4 1-hr recitations and 2 2-hr laboratory periods;
Mathematics, 10 hr/wk, 5 1-hr periods in algebra and 5 1-hr periods in trigonometry;
Orientation lecture, 1 hr/wk;
Supervised study, 11 hr/wk.

During the last five weeks, 8 hr/wk of *applied mechanics* was substituted for mathematics and 13 hr/wk was devoted to supervised study.

It was expected that each student would do a considerable amount of studying outside of the scheduled 40 hours. This hope was realized in widely varying degrees, depending upon local and individual conditions.

Students admitted to the program were high school graduates with a minimum of two years of mathematics, including algebra and some geometry, and one year of some science. Each student was recommended by his high school principal and was personally interviewed by a representative of the College. Many had had more than two years of high school mathematics and some had taken four years. A considerable number had taken high school physics, although little laboratory work involving student manipulation of apparatus was indicated.

Standardized aptitude tests given to each student near the beginning of the course showed that the group selected was definitely superior in physics aptitudes to those regular college students for whom norms were available.

Except for textual materials, all costs were borne by the Federal Government. The cost to each student was approximately $15, including text- and reference books, notebooks, a slide rule and a deposit for drawing instruments.

ADMINISTRATION

The Pennsylvania State College has long maintained large and well-organized extension services, whose experience and facilities enabled the College to undertake this large, widespread program quickly and efficiently. The several schools of the College have directors of extension

[1] I. H. Solt, "Training of physicists for defense industries," Am. J. Phys. 9, 294–96 (1941).

in their respective fields. A supervisor and several assistants are in charge of each subject-matter field. The central extension services maintain district offices in ten important cities of the state. Each district representative is charged primarily with administrative and nonacademic matters; he surveys the industries of the state to ascertain the demand for extension services, familiarizes himself with local needs and reactions to these services, locates centers of instruction, procures building space and secretarial help, and keeps the streams of equipment and supplies moving. At each local center of instruction is an administrative head of the teaching group, who represents that group's local, community and College interests.

In strictly academic matters, supervision comes from the subject matter departments of the College. Thus the physics supervisor represents and is responsible to the head of the physics department of the College; he is charged with maintenance of proper academic standards in such matters as content of courses, approval of teaching personnel, selection of textbooks, teaching aids, laboratory manuals, and so forth. Control in nonacademic matters comes from the officers of the extension division and is exercised in matters relating to the physical plant, finance, time schedules, procedure coordination, and so forth.

In the summer program 84 of the centers were operated in high school buildings, city and rural; 7 were located on college campuses; and 4 were at the regular junior college undergraduate centers permanently maintained by the College. The centers were opened on a staggered date program, beginning June 14 and continuing six days each week, the last center opening on June 23.

Apparatus and Supplies

At most of the centers local facilities for laboratory work of collegiate grade were entirely inadequate. It soon became evident that considerable equipment would have to be supplied in order to make such laboratory work possible. Every prospective center was therefore inspected and the available physics apparatus inventoried by a representative of the physics department of the College to ascertain what additional equipment would be needed.

During the session the department distributed on loan to the various centers much of its own equipment and ordered a great deal more.[2] Unfortunately, many delays were encountered in the attempt to complete this ambitious project. The greatest handicaps that had to be overcome were the necessity of purchasing everything on bids and the fact that it was impossible for the instrument companies in so short a time to supply the large numbers of specialized pieces of equipment needed. Apparatus purchased for this program remains the property of the Government and is stored at the College for use in future government-sponsored courses.

Not until the program was well under way was a satisfactory adequacy of laboratory equipment achieved. In the meantime teachers did much improvising in the laboratory. A great deal of locally made and ingeniously assembled ten-cent store apparatus appeared in many centers. It is believed that reasonably satisfactory work of collegiate grade was done in spite of these difficulties.

Teaching Personnel and Loads

Teachers were selected on the basis of their formal records of training and experience, supplemented in many cases by previous personal acquaintance with staff members at the College. In general, only experienced college physics teachers were employed. However, the faculty included some high school teachers of exceptional qualifications and a few graduate students.[3] Little difficulty was experienced in securing properly qualified, enthusiastic teachers for the summer program.

Each teacher was asked to spend one day in conference at the College before the opening of classes. This day was spent in orienting him to all aspects of the program. More than two weeks were used for the various groups in these

[2] The details of choosing, ordering, delivering and returning materials and supplies were handled by Dr. Henry W. Knerr and four assistants.

[3] There were 127 members of the teaching and supervisory staff in physics, recruited from 91 institutions in 30 states. The teachers have taught an average of 11.3 yrs in college. About 64 percent have the doctorate, 29 percent the master's and 7 percent the bachelor's as their highest degree. Their average age was 36 yrs. The 105 college teachers had an average of 0.89 yrs of teaching experience at the high school level.

daily conferences. A model laboratory was arranged in the physics department and each teacher spent some time in examining the suggested set-ups.

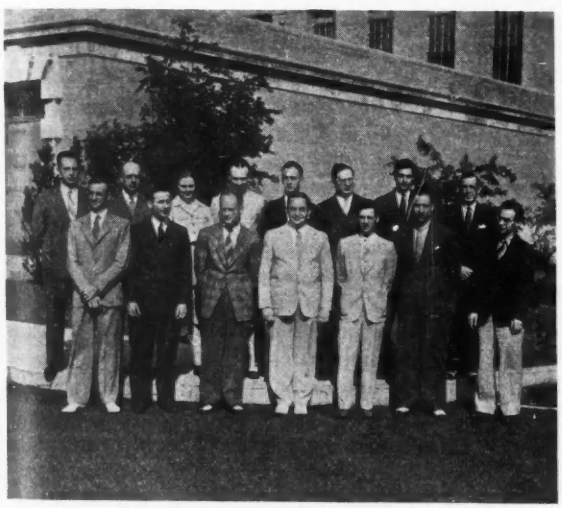
Classes varied in size initially from 10 to 25. When the enrolment at a center approached 25 students, two classes were formed. Although most centers had only one class, there were a number with two and others up to eight.

An average academic load for a physics teacher was 16 clock hours per week. In some cases this time was spent in teaching two physics classes, in others it involved one physics class and several periods of supervised study. Frequently physics teachers taught one course in physics and one in mathematics or applied mechanics.

Supervision of Teachers

Staff meetings for so large a group of teachers scattered in widely separated centers were impossible. As a substitute, the state was divided into five districts and an assistant supervisor was assigned to visit the centers in each of these districts soon after the opening date and at frequent intervals throughout the program. The assistant supervisors were all experienced physicists and successful teachers. They were directly responsible to the supervisor of physics, whose headquarters were at the College. It is believed that the coordination which they achieved is a unique and significant contribution to the success of the whole program.

The key supervisor planned and presided at the weekly conferences of the supervisory staff, held at the College. At these meetings the assistant supervisors acted as proxies for the teachers whom they had visited during the week in their capacity as representatives of the College. No effort was made specifically to define their duties, since the program was new and the problems that might arise could not be clearly foreseen. They were to maintain personal contact with the teachers on the job, to see that the best possible use was being made of all the existing facilities at each center, to assist in arranging for equipment and supplies, and generally to coordinate the work. The apparent needs and resources of each center were carefully examined in the light of the supervisors' common experience. Possible adaptations from among these resources were thus discovered, which facilitated the development of the laboratory work, especially at some centers that began with limited equipment. One physics teacher in an isolated center found that curtain rods made worthy substitutes for more traditional levers when delivery of promised equipment was unavoidably delayed. Another teacher used the spring from a window-shade roller to illustrate the laws of elasticity and of simple harmonic motion. Through the supervisory staff such ideas were reported within a week, with stimulating effect on any teacher in the state who had thought that to perform experiments he would be forced to wait for appropriate apparatus as illustrated in the manual.



Physics supervision, research, apparatus and lecture-demonstration staff. Front row, left to right: Henry L. Yeagley, supervisor of lecture-demonstration tour; Paul M. Kendig, assistant supervisor; Fred Tracy, assistant lecturer; H. L. Van Velzer, lecturer; J. J. Gibbons, assistant supervisor; D. P. LeGalley, assistant supervisor; Wayne Webb, assistant supervisor. Back row: Robert F. Paton, assistant supervisor; C. J. Lapp, supervisor of physics research; Helen Calkins, statistician for physics research; Henry W. Knerr, supervisor in charge of materials and supplies; Robert L. Weber, assistant supervisor; H. K. Schilling, lecturer; J. Lyle Redrup, assistant lecturer; Marsh W. White, physics supervisor.

The ability of each supervisor to initiate the prompt functioning of each center was

The physics "autolab"—mobile demonstration unit.

considerably increased after the first week-end staff meeting. Each of the workers at the central office as well as the supervisors who had been calling at centers all over the state had suggestions in answer to questions that arose. Routine procedures, essential in so large a program, were discussed and misunderstandings were clarified.

A major assignment of the assistant supervisors was the actual observation of recitation and laboratory teaching. Thus good teaching was quickly recognized and a few cases of inept or indifferent teaching were rectified. The assistant supervisors were able to see needs and pass on suggestions that contributed definitely to the coordination of the program. Their itineraries were planned a week in advance, so that notices of the time of visits could be mailed from the supervisor's office to each teacher and administrative head. The teachers thus anticipated the visits and could present their ideas and questions more clearly and completely. A written report was made by the assistant supervisor after his visit to each center and copies were distributed to the physics supervisor, the district representative, the school director of extension and to a master file in the main offices at central extension headquarters. Thus all concerned could follow the work of the teachers.

Teaching Aids

Suggestions and proposals that came up at the weekly meetings of the supervisory staff were organized and supplemented by the key supervisor, and were mimeographed and mailed weekly to each teacher. These weekly bulletins served admirably to provide uniform answers to such questions as when needed apparatus might be expected, what subject matter should be stressed or omitted, where and how to get help in case of emergency, how to utilize study periods, and what testing schedule was to be followed. The weekly news letter thus became an effective tool that did much to knit the centers together, to make each teacher realize that he was not isolated and alone, and to foster his interest and enthusiasm in the program as a whole.

Considerable study has been given to visual aids in the EDT work. Eight motion picture films on topics in physics and chemistry were used and a staff of four operators was employed to commute among the centers on a pre-arranged schedule. Each teacher was given a description of the physics films in advance of their showings and mimeographed aids were distributed to each student so as to increase the effectiveness of the showing.

Mobile Lecture-Demonstration Unit

Equipment suitable for lecture demonstrations was virtually nonexistent in most of the centers. To overcome this difficulty, at least in part, it was decided to develop a mobile physics laboratory that could be used by a traveling unit to give demonstration lectures. The design and completion of this project constituted one of the most spectacular and valuable contributions of the summer program. Started at an unfortunately late date in the session, it was rushed to completion and put on the road about two weeks after work was started.

The lecture demonstrations were designed to be definite teaching aids to the portions of the course dealing with electricity and magnetism. The "autolab," as it was popularly called, also proved to be a valuable publicity source. The unit was called "Electrons at work." It was usually presented at two centers each day. In general, the students from several nearby centers met at a centrally located point for a lecture. The public was invited. There seems to be considerable evidence that such mobile lecture units

fill a real need and constitute a significant addition to the teaching efforts of this type of extension program.

Both teachers and students were provided with a synopsis of the lecture and a description of the exhibits. The teachers were expected to discuss these with their classes before the lecture and again when the subjects came up in class. After the lecture the students were given an opportunity to examine and operate a number of exhibits of the museum or "push-button" type, many of which were quantitative.

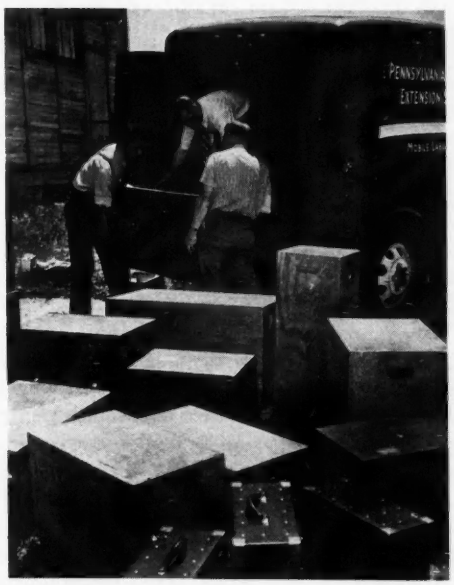
A special woodworking shop and a group of mechanics were obtained to make certain pieces of demonstration equipment not available from instrument companies or the College stock room. The uniquely designed boxes required to transport much fragile apparatus were a particular problem and their manufacture in a short time presented many difficulties. The unit was transported in a two-ton truck. The unusual methods for rapid packing and transporting pieces of scientific equipment were partly adapted from some of those used in the Air Show of the Franklin Institute, whose services were generously offered to the College in the creation of the "Penn State Mobile Demonstration Laboratory."[4]

Textbooks

A textbook[5] and a laboratory manual[6] for the course were purchased by each student. These books were chosen to present the topics to be studied in concise form but with sufficient completeness and rigor to be commensurate with the college quality of teaching expected in the program. In a course of this unique character it would naturally be impossible to use a printed textbook and manual immediately available in large numbers and at the same time ideally satisfactory to every teacher. The books used were selected after considerable study as being the most satisfactory ones known to the supervisory staff.

The fields of mechanics, fluids, heat, electricity and magnetism comprised the course of study, particular emphasis being placed on the first two. Daily assignments included from three to ten problems. Preparation of these assignments was made mostly during supervised study periods, a method that was found to be more productive than relying on any appreciable amount of home study.

The laboratory manual contained the following selection from standard experiments in college physics: verniers and micrometers, vectors, torque, momentum and ballistics, simple harmonic motion, friction, simple machines, Young's modulus and Hooke's law, density and Archimedes' principle, Boyle's and Charles' laws, linear expansion, calorimetry, hygrometry, electrical instruments, cells and resistors in series and parallel, resistivity and the Wheatstone bridge, heating effects of an electric current, voltmeter and ammeter multipliers, and electrolysis. Considerable variation was allowed and expected in the adaptation of the experiments suggested, due



Some of the containers designed to facilitate safe transportation and quick packing of the demonstration and exhibit apparatus.

[4] Dr. Harold K. Schilling and Dr. H. L. Van Velzer were the lecturers on the tour; they were accompanied by three assistants. The arrangements for the tour and other administrative details were under the supervision of Dr. Henry L. Yeagley.

[5] J. E. Hoyt, *A concise physics for engineering students* (Blakiston, 1940).

[6] M. W. White, *Experiments in essentials of engineering physics* (McGraw-Hill, 1941), being sections reprinted from the same author's *Experimental college physics* (1940).

to local conditions of instruction and equipment available. Each center was expected to perform the equivalent of at least 16 of the 19 experiments suggested. Experience with laboratory instruction in the present program suggests the advisability of including in the laboratory manual numerous alternative experiments involving very simple and easily obtainable equipment.

The detailed course outline, provided for each teacher and student, contributed in a large measure to coordinating the assignments and making them effective. The syllabus represented a careful selection of topics and a judicious attempt to fit the length of each assignment to the anticipated ability of the students to assimilate it. Some teachers expressed dismay at the prospect of covering selected portions of four major fields of physics in ten weeks; but who has not met similar objections to the strenuous pace of any college physics course?

The outline contained daily assignments devoted to the chief topics as follows: mechanics, 17; fluids and properties of matter, 5; heat, 5; electricity and magnetism, 8. Five periods were nominally assigned for review, tests and examinations. The outline encouraged coordination of classroom and laboratory work by including references to the parts of the laboratory manual that involved theory.

A handbook[7] and a slide rule were part of each student's equipment. Each teacher was provided with a copy of *Demonstration experiments in physics*[8] and of *Selective experiments in physics*.[9] These, it was expected, would offer practical suggestions for lecture demonstrations and for alternative experiments. While these books were helpful, experience has shown that supplementary books based on even simpler equipment would be desirable.

Instructors were requested near the end of the program to give detailed criticisms of the textual material used in the courses. Many of the criticisms canceled one another. Taken together, they seemed to reduce to the statement, made by some, that the teaching materials used were the best available.

[7] Hodgman, *Handbook of chemistry and physics* (Chemical Rubber Publishing Co., 1939).
[8] R. M. Sutton, ed. (McGraw-Hill, 1938).
[9] A set of looseleaf sheets currently available in mechanics and heat (Central Scientific Co., series *b*, 1941).

A typical class in one of the centers visited by the mobile demonstration unit.

Research and Testing Program

Since this course probably involved the largest number of students ever enrolled in a single course in college physics and since satisfactory facilities were available, an opportunity for the study of physics teaching never before existent seemed to be presented. Consequently a research program[10] was designed and made an integral part of the teaching work. This program was facilitated by the availability and use twenty-four hours per day of several Hollerith computing machines. Many contributions are expected from this phase of the summer's work, including certain evaluations of teacher ability, better information with regard to the qualifications of a good physics teacher and extensive data that will furnish the basis for further studies.

The *Iowa physics aptitude test*, Form M, and the national *Cooperative physics test for college students*, Mechanics 1936 A, were given to each student near the beginning of the course. The former test purports to measure the student's aptitude for physics, particularly his ability to handle algebraic symbols, to formulate a symbolic statement from problems stated in words and to read and reason about the facts therein, and his interest in physics. The latter test measured the student's knowledge of mechanics prior to beginning the study of that subject.

As a part of the general EDT testing program at the College centers and not connected with

[10] Directed by Dr. C. J. Lapp, State University of Iowa, assisted by Dr. Helen Calkins, head of the mathematics department, Pennsylvania College for Women, as statistician, and a group of clerical and statistical workers from the local EDT research organization.

the physics research, all students were required to take the following tests: Strong, *Vocational interest;* Moore, *Arithmetical skills;* Otis, *Self-administering test of mental ability; Minnesota vocational test for clerical workers; Minnesota paper form board;* Bennett, *Mechanical comprehension.* The scores on all of these tests are available and will be used in evaluating the experimental procedures used in the physics teaching program.

Two special technics were used in the research program. Half of the teachers, selected by chance, employed these two methods; the other half served as a control group and taught only in the conventional manner.

The first special method was that of *systematic review.* The object of the whole teaching project is "to get as much useful and usable physics actually into the students as possible." It is known that consecutive lessons "go into the students in layers." These layers must be "turned on edge" so as to make all of the material available. Two review devices were suggested for use by the teachers in the research group:

(1) A summary of the minimum essentials for each lesson, which was handed to the student *after* the teacher had taught the lesson in his own way; and

(2) A cumulative review sheet, mimeographed on the bottom half of the summary sheet. The student was instructed to spend not more than 10 min a day on this review. It consisted of questions and problems intended to keep the learned material available and usable. These materials were furnished from the physics office for each lesson and each student.

The second special technic involved the *physical solution of problems.* It is desired in physics to teach the student to think clearly. The most effective way to do so is to exercise him in thinking. The physical solution of a problem teaches a systematic attack which is akin to the scientific method. An outline of the method was given to each student after the teacher had explained its meaning. This part of the program was most exacting. It demanded:

(1) That every student solve and hand in a minimum of three problems with each lesson; the teacher selected these problems and could require as many more as he chose;

(2) That all problems solved or demonstrated by students or teacher be attacked by the physical method.

Experiments along this line have already been conducted and the method is known to produce favorable results.[11] It appears to be time-consuming at first, but as soon as the students become accustomed to it they actually solve problems faster and achieve more.

To evaluate the results of these special methods, two examinations of 60 min each were required of all students. The first of these was a problem examination consisting of three parts, two of which dealt with critical analysis. The second was the national *Cooperative Physics Test for College Students,* Mechanics 1936 B, which served as a post-test and measure of achievement in this field.

From the pre-testing program 11 measures of the student's ability were available. It is planned to use the results of the mechanics post-tests to compute a regression equation for the pre-test measures. The weighting so obtained will be used to determine a coefficient that predicts each student's performance. The differences between the predicted and the actual performances on the mechanics post-test will then be studied for both the research and the control groups. The results of these studies will be announced later.

The success of the program as a whole can only be measured by its long-range results. However, gratifying reports of the benefits obtained are already being received from industries, community leaders, faculty and students.

This report has been prepared with the collaboration of various members of the physics supervisory staff.

[11] See, for example, C. J. Lapp, Am. J. Phys. **8**, 241 (1940); **9**, 239 (1941); H. K. Schilling, Am. J. Phys. **8**, 68 (1940).

Consistency in regard to opinions is the slow poison of the intellectual life, the destroyer of its vividness and its energy.—SIR HUMPHRY DAVY

Secondary School Physics in Arizona

E. H. WARNER
University of Arizona, Tucson, Arizona

SEVERAL years ago the United States Commissioner of Education published data which gave the extent of the high school instruction in natural sciences in 1934 and the changes that had occurred in this instruction during a six-year interval. The results are summarized in Table I and in the upper part of Table II.[1]

The poor showing made by physics makes one question the nature of high school physics instruction. In order to study the extent and nature of this instruction during the school year 1940–41 in the state of Arizona, a questionary[2] was sent to every high school physics instructor in the state. Every high school replied. The data and conclusions of this paper were obtained from these answers.

The result of the extent of the instruction in physics is shown in Table II. The percentage of enrolled students who elected physics varied with the size of the school, as shown in Fig. 1. The fact that physics is offered in only about 60 percent of the Arizona high schools and elected by less than 5 percent of the students forces one to conclude that the position of the science is not at all satisfactory. It may be in final decline. The opinions of the teachers themselves should be considered. In answering the question whether physics is on the increase or decrease in their schools, the instructors reported: stationary, 44 percent; increasing, 44 percent; decreasing, 12 percent. Most of the teachers who reported increased interest in physics commented that such interest was very slight. So the teachers themselves seem to think that physics is doing just a little better than holding its own.

In order to obtain information that might help to explain this condition, the questionary asked for data on four items: teaching load; training in physics possessed by the teacher; value of physics equipment; annual appropriation for physics.

TABLE I. *Continental United States.*

	1934		1928 to 1934	
			SCHOOLS	STUDENTS
SUBJECT	ENROLMENT (PERCENT)	% OF SCIENCE ENROLMENT	% OFFERING 1934 / % OFFERING 1928	% TAKING 1934 / % TAKING 1928
Gen. Sci.	15	38	1.16	0.86
Biology	13	33	1.21	0.89
Chemistry	6	15	1.15	0.90
Physics	5	13	0.97	0.76

TEACHING LOAD

The questionary showed that 61 percent of the physics instructors teach 5 periods per day; 19 percent teach 6 periods per day; 5 percent teach 7 periods per day; only 15 percent teach less than 5 periods per day. These programs are exclusive of study hall supervision.

A load of 5 periods per day, plus study hall in many cases, plus paper-grading, extracurricular work, and so forth, simply does not leave time to prepare and set up apparatus for lecture-table demonstrations and laboratory work. It is almost certain that physics classes taught by teachers with these loads will not stimulate student interest or attract larger enrolments. The physics teachers are overloaded. This fact alone is an important cause of the present standing of physics in Arizona high schools.

COLLEGE TRAINING

The college training in physics possessed by Arizona high school physics teachers is shown in

TABLE II. *Continental United States and Arizona.*

		SCHOOLS	STUDENTS	SCHOOLS OFFERING PHYSICS No.	%	STUDENTS IN PHYSICS No.	%
1934	Cont. U.S.	17,897	5,402,305	8466	47	282,820	5.2
	Arizona	55	15,501	28	51	666	4.3
1940	Arizona	59	19,509	33	56	867	4.4

[1] From data supplied by C. A. Jensen and L. V. Herlihy, Sch. Life **22**, 314 (1937). This article is briefly reviewed in Am. J. Phys. (Am. Phys. T.) **6**, 53 (1938). See also Fountain, "Physics for the masses," Am. J. Phys. **8**, 135 (1940). For similar study for one county in Pennsylvania, see Trytten, Am. J. Phys. **8**, 54 (1940); also, for a sampling of schools in Pennsylvania, Trytten and Leach, Am. J. Phys. **9**, 96 (1941).

[2] Similar to the one used by Trytten, reference 1.

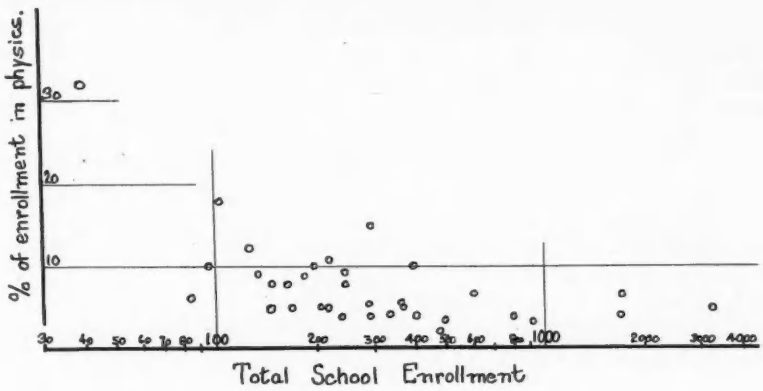


FIG. 1. Enrolments in school physics.

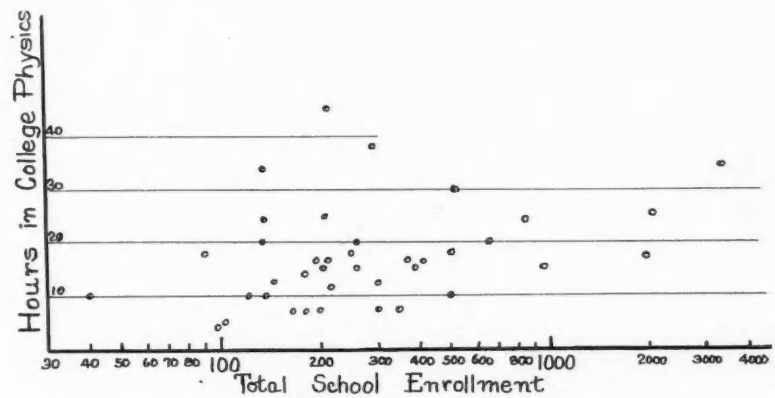


FIG. 2. Teacher training.

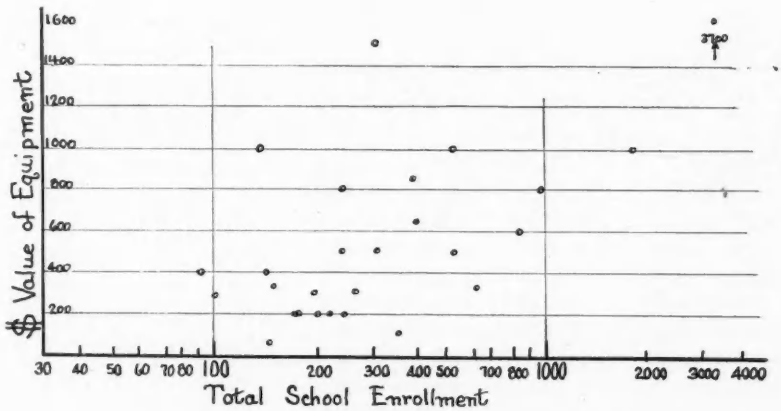


FIG. 3. Value of equipment.

Fig. 2. It will be noted that the teachers in the larger schools have had considerably more college training in the science than teachers in the smaller high schools.

Ten percent of the physics teachers in the state have sufficient credit for a college major in physics, and 22 percent have sufficient credit for a minor. Seventeen percent have 8 hours or less

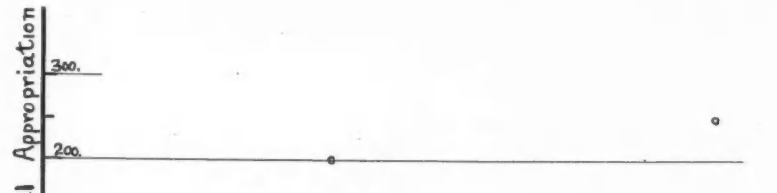


FIG. 4. Annual appropriations.

of college physics to their credit; that is, they have completed, or partially completed, only the introductory course in college physics. Sixty-eight percent have not had college physics courses sufficient for a minor in the field. The average college training in physics for the whole state is 17 hours.

The spirit of physics and the enthusiasm for physics possessed by workers in this field are usually acquired only by students who have majored or at least minored in physics in college.

VALUE OF PHYSICS EQUIPMENT

The value of equipment in physics is shown in Fig. 3. The average is $680, an exceedingly low figure. A large amount of apparatus is necessary in physics, much more than in other sciences. Physics taught without adequate equipment for laboratory and demonstration is necessarily weak. Physics requires considerable visualization, an ability which seems inadequately developed in young people. Student use of appropriate apparatus would improve this condition. Courses taught with insufficient equipment become unintelligible and hopelessly confusing.

The limited amount of money invested in physics equipment is, without a doubt, an important factor in explaining why physics is not attracting more students in our high schools.

ANNUAL APPROPRIATIONS FOR PHYSICS

The annual appropriation for physics as distributed between the schools is shown in Fig. 4. The average for the state is $80 per year per school. Since the average life of apparatus in our field is about ten years, a school should be spending about 10 percent of the value of equipment on maintenance and replacements. An additional 10 to 20 percent should be spent yearly for improvement and modernization. Thus each school should have an annual budget of from 20 to 30 percent of the equipment value. This would mean for the average Arizona high school an annual appropriation of $175.

Arizona high school physics teachers are working under a tremendous handicap when forced to carry on with an average annual appropriation of $80.

GENERAL DISCUSSION

It is possible that the deplorable state of physics in the schools is due to factors entirely separate from the training and enthusiasm of the high school physics teachers. The condition may be largely caused by the unpopularity of physics with administrators, who generally are well trained in professional education and administration but have little training in the sciences. With this background they are unable to appreciate the educational value of a thorough course in physics.

There is a tendency in Arizona high schools which do not offer physics to present certain skimmed-milk courses, appealing to popular interest only, such as photography, descriptive physics, industrial science, and general science survey. The trend in high school education seems to be that only pre-engineers and those planning

a lifework in science should take physics. All others should elect almost anything else.

If this be the case, all of us who are interested in physics should know about it, and we should institute a program aimed at showing school administrators, teachers and students that physics has cultural value; that it furnishes the opportunity to think accurately and comprehensively; that it is a first-class subject for developing one's power of observation; that without this science a student can only inadequately comprehend the modern scientific world-picture; that even a slight knowledge of physics can help one in his everyday work and that such knowledge will give him more enjoyment in life by enabling him to understand the ever present applications of physics; that no other science course will enable one to appreciate and to apply the modern scientific method better than a course in physics.

Unless college science teachers who know the value of physics in our educational scheme can convince our high school communities that these ideas are true, we may in the near future find that physics in the schools has dwindled to a point where its offering is economically unwise.

Teaching Physics by the Conference Method

R. R. HANCOX
*Wilson Junior College, Chicago, Illinois**

FOR the past several years attempts have been made at Olivet College to teach elementary physics by the conference method rather than the more usual lecture-recitation method. Different plans have been tried, in which the time devoted to different parts of the work has been varied.

Olivet is a small denominational college with an enrolment of about 300 students. Physics is not a required subject and this circumstance may have resulted in a somewhat select group. However, as far as can be determined from psychological and other tests, the students taking physics were of average ability. The majority were premedical students or majors in biology, chemistry or physics. The number of students in the first year of physics ranged from 15 to 25.

Several experimental plans were tried before a satisfactory one was found. In one plan the class was divided into sections of five or six students each. These sections met once each week for conference. In addition, there were two 3-hour laboratory periods for each student, one of which was used principally for demonstration experiments. Whenever possible and practicable the students were allowed to set up and demonstrate the experiments, and to lead the discussion which arose from the demonstration. In the other laboratory period the experiments were of the usual type.

In another plan each conference section met twice each week with only one 3-hour laboratory period. In this plan demonstrations were made during the conference periods. Other plans tried included combinations of the lecture and conference methods.

The following plan is the one that finally proved to be the most satisfactory. The entire class met twice each week. The first meeting was devoted to demonstrations and to such material as could not readily be obtained from textbooks. The second was used almost entirely for problem work, with emphasis particularly on the physical solution. In addition, the class was divided into conference sections of about six students each and each group met once each week. There was also one 3-hr laboratory period. Many experiments were listed and students were allowed to choose a certain number from the list. No single laboratory manual was employed; instead, several manuals were made available and students were asked to outline their experiments before coming to the laboratory and to have their proposed methods checked by an instructor before beginning work. No single textbook was adopted for the class as a whole. Instead, copies of eight different textbooks were placed on

* Formerly at Olivet College, Olivet, Michigan.

TABLE I. *Total scores of 27 students.*

PRE-TEST	POST-TEST	GAIN	PRE-TEST	POST-TEST	GAIN	PRE-TEST	POST-TEST	GAIN
68*	136	68	30*	82	52	19*	67	48
28	105	77	17*	81	64	19*	63	44
48	103	55	24	78	54	16	56	40
33*	95	62	39*	76	37	21*	53	32
35*	95	60	22	76	54	5*	47	42
28	90	62	25*	76	51	10	37	27
23	87	64	44*	74	30	16*	36	20
23	85	62	13*	70	57	2	32	30
17*	83	66	17*	67	50	8	25	17

* Had high school physics.

reserve and the students were asked to read from more than one book. Actually, only a few students followed these directions. Many, after a preliminary investigation of several books, chose one which they used almost entirely from then on. However, even for those students who studied from one book only, it is believed some benefit may have resulted from this method, since no single textbook seemed to predominate markedly in the selections of the students. In the conference groups different approaches and points of view were repeatedly brought out.

Students were asked to come to the conference prepared to discuss the topic under study at the time. These discussions were sometimes written, sometimes oral. After a report was given the other students asked questions and made criticisms. There was thus an excellent opportunity to clear up points of uncertainty and for the instructor to get a clear idea of each student's progress. A comprehensive examination was given over the work of the entire year.

During the first weeks of the year some time was usually lost before the students felt free to discuss and criticize reports given by others in the group. However, as emphasis was placed on the fact that the conferences were the best opportunities for each student to prepare himself for the final comprehensive examination and as the members of the group became acquainted with one another and the instructor, this restraint was broken down and then real progress was made. The whole plan aimed at establishing in the student's mind the idea that responsibility for his progress rested with him and not with the instructor.

It is evident that this plan could not be used with a large class. With the small groups at Olivet it proved quite satisfactory. Students enjoyed the opportunity for small group discussions.

The *Coöperative Physics Tests* were given as pre-tests and again as post-tests upon the completion of each branch of the general course. Over a two-year period complete records were obtained for 27 students. This number is too small to make the results very significant; however, they are summarized in Table I. Students who had high school physics did not show any superiority on post-test scores; in fact, more than half of them ranked below the median. The correlation coefficient between pre-test and post-test scores is 0.75, and between pre-test scores and post-test gains is 0.54. I.Q. scores are available for 21 students. There appears to be a high correlation between these scores and the gains made, but a considerably lower correlation between I.Q. scores and either pre-test or post-test scores. It is believed that the gains made by this method are somewhat higher than had been found by the author using other methods of instruction.

ADVANCED WORK

In the advanced courses in physics the work was done almost entirely individually. Each student's course of study leading to graduation was tentatively outlined in the light of his interests and needs. Each unit of study was continued until the student had covered satisfactorily at least a minimum of work. Although the requirements were varied according to the students' ability, more being demanded of a good student, in no case was the requirement made less than the amount that the instructor considered to be adequate for a good grade of college work.

Each major student was required to pass a final comprehensive examination over all work taken after the first year which related directly to the major. This always included mathematics and sometimes chemistry as well as his work in physics. The final examination required approximately 27 hr, being divided into 9 parts of 3 hr each. Admission to it required the recommendation of the student's adviser and final authorization by an academic committee. The student was recommended only after the completion of a

certain minimum amount of work, which again was varied according to the ability of the student. Students were discouraged from attempting to finish in less than four years and only rarely allowed to do so, in the belief that a valuable part of college experience would be missed by too rapid coverage. It was thought better policy in most cases to require more work of the superior student than to encourage graduation at an early date.

A record of the material covered satisfactorily was kept but was not measured in terms of hours of credit, although for purposes of transfer, the conversion could readily be made. This was done to relieve the student's mind of the idea that his graduation was conditioned upon the accumulation of a certain number of credit-hours. The emphasis was put, instead, on the amount of material on which both the student and the instructor felt he was prepared to write a final examination. It is the author's belief that this resulted in a better attitude on the part of the student toward his work and to a more comprehensive view of the field of physics as a whole.

Laboratory Experiments on Radioactive Recoil

SANBORN C. BROWN
Massachusetts Institute of Technology, Cambridge, Massachusetts

THE separation of isotopes by radioactive recoil can be made the basis of a number of laboratory experiments. Students in a course in experimental nuclear physics given at the Massachusetts Institute of Technology seem to find experiments of this kind interesting and instructive enough to warrant publication of a description of the methods and apparatus used.

As is well known, when a radioactive body decays with the emission of an α-particle or a β-ray, the emitted particle travels in one direction, and to conserve momentum, the resulting nucleus must travel in the opposite direction. When this process happens to take place near the surface of a radioactive source, and if it is in a vacuum so that the mean free path is fairly long, one may collect the recoiling atoms on a plate placed near the source. This therefore provides a nice method for separating a daughter substance from its parent.

Figure 1 is a general view of the apparatus. A bell jar fitted with a ground joint is waxed onto a brass plate into which are built a Geiger-Müller counter and a copper tube through which a source may be introduced. The ground joint makes it possible to turn the two brass disks which act as recoil collectors, so that one particular collector can be placed first over the source and then over the counter. The end of the Geiger-Müller counter, which fits through the brass plate, is open, with the center wire supported by a glass cross fitted over the edge of the plate. A lead block between the source and the counter is essential if the experiment is to be carried out with thorium recoils, because of the strong γ-ray from the ThC''. It is not necessary if actinium is used. The inside of the bell jar must be coated with a conducting paint and kept at a potential of about 1500 v above ground throughout the experiment, otherwise ions drifting into the counter will cause a background of several thousand counts per minute.

Before the experiment can be started, the ThB and AcB sources must be made. This can be done by suspending the source plate in a closed vessel in which a source emanating thoron or actinon is also placed. For the emanation sources used in this laboratory, a negative potential of 1000 v is placed on the suspended source plate, the vessel acting as the other electrode, and the vessel is evacuated to a pressure of 6 cm-of-mercury.

In performing the experiment, the student removes the source plate from the thoron or actinon gas atmosphere in which it has been collecting ThB or AcB, as the case may be, for a sufficiently long time to have attained maximum activity. He waxes the source plate into the tube which is built to receive it. One of the recoil collectors should always be kept over the source so that the bell jar does not become contaminated with recoil atoms. The bell jar is now pumped down with a Cenco "Hyvac" pump, and the system is checked to make sure it does not

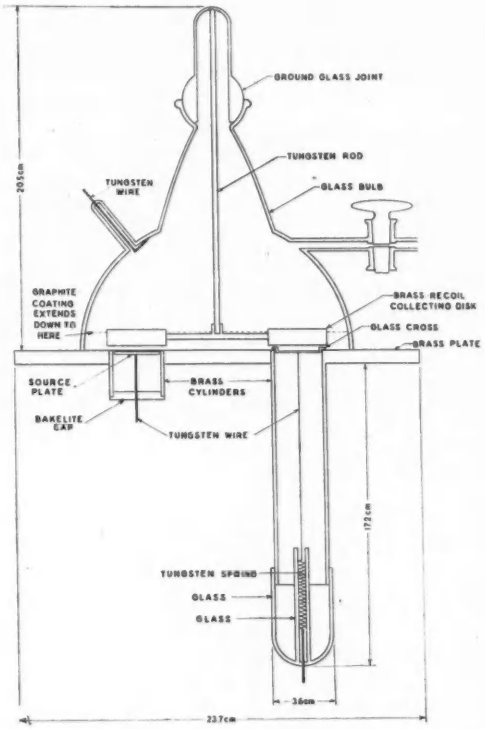


FIG. 1. General view of the apparatus.

leak. Then, with the apparatus still pumping, the collector which has been over the counter is moved into position to collect recoils over the source plate and is left there for 10 min. During this period of collection, the ion-clearing field should be turned on. At the end of 10 min, the bell jar is filled to a pressure of 6 cm-of-mercury with dry air and the recoil collector is swung into position over the counter and its activity followed for an hour. The first 20 min will be sufficient to determine the decay of the recoil atoms of the C″ bodies and the last 20 min will give a background reading which must be subtracted from the observed activities before the data can be used for the determination of the half-life. It has been found in the author's laboratory that the experiment can be performed about twice in the course of a 3-hr laboratory period and that the averaged results agree with published half-lives within the experimental error. To repeat the experiment, the collector plate should be turned back to its position over the source for another 10-min period after the bell jar has been pumped down.

The most accurate data are taken when the counting rate is high in the first few minutes after the collection has stopped. So that the student does not waste time at this point of the experiment, in getting the gas pressure in the counter just right and adjusting the voltage of the amplifier, it has been found convenient to arrange a gas reservoir between the bell jar and the pump. After the bell jar has been pumped down, the stopcock is closed, and during the 10-min recoil collection time, the student adjusts the pressure of air in the reservoir to a previously determined value such that, when the stopcock is opened, the whole apparatus will come to a pressure of 6 cm-of-mercury. The voltage on the counter should be left so that this pressure will bring the counter onto its plateau. Then all the student has to do, between the time the recoil collection stops and the time when he is ready to take the data, is to open the stopcock between the bell jar and the reservoir and swing the collector plate from the source to the counter.

If the foregoing procedure is followed, the student will obtain either the 3-min ThC″ or the 4.5-min AcC″, depending on the source used, separated from all other members of the series, since these are in each case the last radioactive members of the series. Both of these substances have such short half-lives that they offer an extremely useful exercise in statistics. Rapidly decaying substances are not described by the Poisson law, and these particular substances provide a very good problem in the use of Peierls statistics for rapidly decaying sources.

If more than 3 hr are available for carrying out the experiment, further interesting data can be obtained. For example, if recoils are collected for 3 hr instead of 10 min, β-ray recoils of the 60-min ThC will also be collected, so that the activity of the plate will first decay with the combination of the 3-min ThC″ and the 60-min ThC. After the former has completely decayed, the student will find that he can measure the ThC body by itself. Since the ThB on the brass source plate emits a very strong γ-ray, the decay of the 10.4-hr ThB body can be measured by leaving an inactive recoil collector over the counter and removing the lead block from time to time in order to measure the γ-ray activity.

Laboratory Experiments with Acoustic Resonators

PAUL L. COPELAND
Illinois Institute of Technology, Chicago, Illinois

PROBLEMS concerning the vibration of simple systems and the normal modes of vibration for coupled oscillators are of interest because of their applications in many fields of physics. The principles underlying such problems may be studied experimentally in the acoustic case without great difficulty. Such experiments can be made suitable for students in the general physics laboratory. A need for additional sound experiments to supplement those ordinarily covered in the general laboratory has long been felt, and in this paper several experiments with acoustic oscillators will be discussed. One of the merits of this type of work is the wide variety of material, both experimental and theoretical, which is available for study.

Simple acoustic resonators

A Helmholtz resonator, suitable for experimental work, may be made by taking two lengths of cylindrical metal tubing of different diameters and connecting them by means of a flange, as shown in Fig. 1. The volume of air, V, enclosed in the large tube, may be varied by raising or lowering the water level w. If a tuning fork is held at the opening O of the small tube, the column of air in it may be set into vibration and if the length l of the small tube is small compared with the wave-length of the sound from the fork, the air in the smaller tube may be thought of as moving up and down bodily, much as would a solid piston. The displacement of the air in the larger tube will be much less, because of its larger cross-sectional area; but, because the tube is closed at the bottom by means of the "water piston," any air that is forced into the large tube would raise the pressure in it. These facts suggest that the chief effect of the air in the small tube arises from its inertia, and that the chief effect of the air in the large cylinder is the production of a restoring force when the air in the smaller tube is displaced in either direction from its equilibrium position. It is customary to consider the inertia of the gas in the small tube, but not its compressibility, and to consider the compressibility of the gas in the large tube, but not its inertia; then the problem of computing the natural frequency of the system is a relatively simple one, and the result of the calculation is sufficiently exact in most cases for comparison with experimental results.

According to the simplifying assumptions just suggested, the mass of the oscillating column will be

$$m=\pi\rho_0 d^2[l+0.72d]/4, \qquad (1)$$

where ρ_0 is the density of the air, d is the diameter of the small tube, l its length, and the term $0.72\,d$ is inserted to take care of end corrections[1] due to one end with a flange and one without. The restoring force constant is given by

$$k=\gamma p(\pi d^2)^2/16V, \qquad (2)$$

where γ is the ratio of the specific heat of air at constant pressure to that at constant volume,

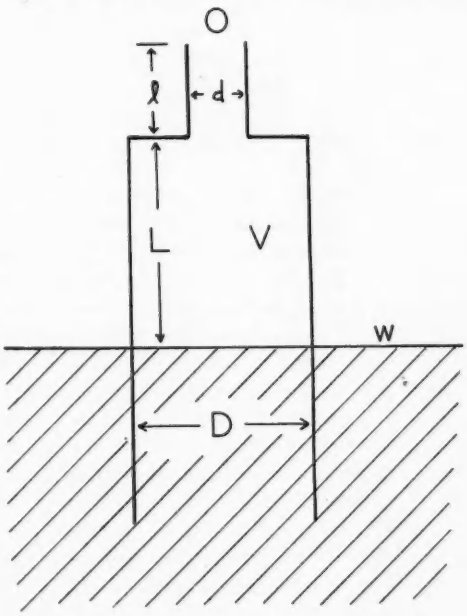


FIG. 1. Experimental Helmholtz resonator.

[1] P. M. Morse, *Vibration and sound* (McGraw-Hill, 1936), p. 200.

FIG. 2. Helmholtz resonator with two vibrating air columns.

and p is the pressure. Thus the natural frequency f_0 is

$$f_0=\frac{1}{2\pi}\left(\frac{k}{m}\right)^{\frac{1}{2}}=\frac{dc}{2}[4\pi V(l+0.72d)]^{-\frac{1}{2}}, \tag{3}$$

where c represents $(\gamma p/\rho_0)^{\frac{1}{2}}$, the speed of sound in air.[2]

With a Helmholtz resonator such as that shown in Fig. 1 and several tuning forks covering a wide range of frequencies, resonance may be established for each fork and the corresponding dimensions entering Eq. (3) may be recorded. The tuning is found to be quite sharp, and comparison of the accepted frequencies of the forks with those computed from Eq. (3) shows a satisfactory agreement.

Basically the same experiment can be performed with the apparatus shown in Fig. 2. Here two similar oscillating masses are connected to the same large volume of air. The mechanical analog of this system is shown in the lower part of Fig. 2. The two large tubes are cut from telescoping tubing, and hence the volume of the air providing the stiffness can be varied by sliding one tube into the other. According to the simple treatment, in which the inertia of the air in the large tube is neglected, a system of this sort can vibrate in only one way. The masses at the opposite ends have to be displaced simultaneously in opposite directions. For the purposes of our approximate treatment we may visualize a node at the middle of this symmetrical system. Hence the frequency can be computed from Eq. (3) by using either one of the masses and a restoring force constant based on half the total volume within the telescoping tubes. Such a resonator is found to have a fairly broad response, but the frequency as computed from the dimensions of the oscillator, when set for maximum response, agrees with the frequency of the fork used to within the limits of the accuracy with which the setting can be made.

If a single pair of telescoping tubes is used for the chamber, the resonator has a very restricted frequency range. This range might be extended, either by using interchangeable small tubes of varying length to change the mass, or by using three or more sizes of tubing each of which telescopes into the one just larger for making the main chamber. The idea of using interchangeable small tubes, all of the same cross section but varying in length, also provides the possibility of making an unsymmetrical oscillating system and investigating its properties. The theory of such an arrangement is a rather simple extension of that just given.

Coupled acoustic oscillators

A system consisting of several similar acoustic oscillators connected to one another is very interesting. For the purpose of studying the properties of such systems, the apparatus shown in Fig. 3 has been designed and constructed. The oscillator shown at the left in Fig. 3, when used alone, resonates at a fixed frequency. However, any desired number of the similar units pictured in Fig. 3 may be connected together by means of tight fitting joints and used as an oscillating system.

For the purpose of locating the frequencies resulting in maximum response of such systems, a calibrated source of continuously variable audiofrequencies of constant loudness would be desirable. A source which has been found satisfactory for this purpose is based upon the beat frequency oscillator described by Harnwell and Van Voorhis.[3] The output from this oscillator is amplified in an auxiliary unit and is then used to drive a crystal loudspeaker. This source, while not giving a note of exactly constant loud-

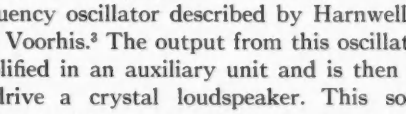


FIG. 3. Apparatus for study of coupled oscillators.

[2] The theory of the Helmholtz resonator is discussed more fully in books on acoustics; for example, Rayleigh, *Sound* (ed. 2), vol. 2, p. 172, and Morse, reference 1, p. 201–202.

[3] G. P. Harnwell and S. N. Van Voorhis, Rev. Sci. Inst. **6**, 194 (1935).

ness when the frequency is varied by a simple condenser adjustment, has no sharp resonances in either the electrical or the mechanical system, and it has proved to be sufficiently uniform for the purposes of this experiment. In practice the speaker is placed in a wooden box lined with material having a high acoustic absorption. In this box a hole is cut for the insertion of one end of the acoustic resonators, and as the frequency is changed, the variation in loudness of the sound transmitted to the room reveals the position of the resonance peaks for the system under investigation. With oscillators such as those shown in Fig. 3, it is observed that there are as many different resonance frequencies as there are units in the system, and these frequencies, of course, correspond to the simple harmonic or normal modes of vibration for the system of coupled oscillators.

The location of these frequencies from theoretical considerations involves the solution of simultaneous equations. As an illustration of the procedure, consider the system formed by the addition of a single unit to that shown at the left in Fig. 3. This system and its mechanical analog are shown in Fig. 4. Here the three oscillating masses are equal and the two coupling stiffnesses are also equal. To simplify the resulting equations, let us agree to start with the system in equilibrium and to measure the displacement of each mass from its equilibrium position. Thus the displacement of m_1 from its equilibrium position is denoted by x_1, that of m_2 by x_2, and that of m_3 by x_3. Then, if the restoring force constant for each compression chamber as a whole is denoted by k, the equations describing the motion of the system are

$$\begin{aligned} m_1a_1&=k(-x_1+x_2),\\ m_3a_3&=k(-x_3+x_2),\\ m_2a_2&=k(-2x_2+x_1+x_3), \end{aligned} \tag{4}$$

where a_1, a_2 and a_3 are the accelerations of the respective masses. Now, for isochronism of the resulting simple harmonic vibrations, we must have

$$a_1/x_1=a_2/x_2=a_3/x_3. \tag{5}$$

Noticing that $m_1=m_2=m_3$, and combining Eqs. (4) and (5), we obtain

$$-1+\frac{x_2}{x_1}=-2+\frac{x_1}{x_2}+\frac{x_3}{x_2}=-1+\frac{x_2}{x_3}. \tag{6}$$

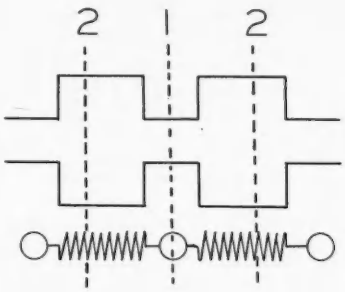


FIG. 4. Pair of coupled oscillators. The dotted lines indicate possible nodal planes.

This set of equations may be satisfied by making $x_2=0$ and $x_1=-x_3$, in which case the frequency f_1 is given by

$$f_1=(k/m)^{\frac{1}{2}}/2\pi=f_0/2^{\frac{1}{2}},$$

where f_0 is the frequency indicated in Eq. (3). The equations may also be satisfied by making $x_1=x_3$ and $2x_1=-x_2$, in which case the frequency f_2 is given by

$$f_2=(3k/m)^{\frac{1}{2}}/2\pi=(\tfrac{3}{2})^{\frac{1}{2}}f_0.$$

The physical interpretation of the first solution is a mode of vibration in which the central mass remains at rest and the outer masses make simultaneous equal and opposite excursions from their equilibrium positions. The node in this case is suggested in Fig. 4 by the dotted line marked *1*. In the second case, the motion is more complicated. The displacement of the two end-masses is in the same direction, while the central mass moves twice as far in the opposite direction. Thus, as a first approximation to the facts, we might visualize nodes located at two-thirds of the distance from the central mass to the outside masses; these positions are shown in Fig. 4 by the dotted lines marked *2*.

When one section after another is added to the oscillating system, the resulting normal modes of vibration and the corresponding frequencies may be found from the same physical principles. The results of such calculations together with observed frequencies of maximum response are shown in Table I.

The predictions of the foregoing simplified treatment may be verified in some detail by using the apparatus previously described. In the first place, each time a section is added to the

TABLE I. *Frequencies (cycle/sec) corresponding to the normal modes of vibration.*

SINGLE-SECTION RESONATOR	TWO SECTIONS		THREE SECTIONS		FOUR SECTIONS	
	OBS.	CALC.	OBS.	CALC.	OBS.	CALC.
Observed and used for f_0 in calculated frequencies					360	366
			460	454		
	590	593			665	697
840			830	840		
	1035	1030			945	962
			1120	1095		
					1135	1130

system, an additional normal mode of vibration is shown in the response of the system. Furthermore, all of these frequencies, as shown in Table I, correspond to those calculated. The data shown here are representative of those generally obtained with the apparatus described, and any differences appearing might arise from experimental error.

Low pass filter

The experiment just outlined suggests how a low pass filter is built from coupled acoustic resonators. As the number of resonator units is increased, the individual resonance frequencies fall nearer one another, and the highest frequency rises more slowly than the lowest frequency falls. For five units, going beyond the range of Table I, the five resonance points were easily distinguishable. For six units, the frequencies were less distinct. When the number of units is made large, the individual resonance frequencies cannot be separated. In this case all the frequencies below a critical value are reduced in intensity or "attenuated." The "critical frequency," used in this connection, is that of the highest normal mode. For an infinite line of resonator sections, this is one in which alternate masses are displaced equal amounts in opposite directions, and for this mode of vibration, the center of each compression chamber may be treated as a node. Thus the frequency is given by

$$f_c = (1/2\pi)(4k/m)^{\frac{1}{2}}.$$

In computing this critical frequency, it was assumed that the number of sections was infinite. The frequency corresponding to the highest mode of a four- or five-section system is found to be fairly close to the critical frequency, and such a system is a reasonably effective low pass filter. It produces very considerable attenuation for frequencies above that of the highest normal mode.

As long as the wave-length of the sound is much larger than the dimensions involved in the individual resonators, the conclusions reached by the foregoing elementary analysis are valid. On the other hand, this reasoning should not be applicable to the higher audiofrequencies, where the wave-length is not much larger than the dimensions, and experimentally a second pass band is observed above a region of pronounced attenuation. For example, in the case of the six-section filter made from sections identical with those yielding the results of Table I, a second pronounced transmission band appeared at about 3000 cycle/sec. Since the discussion of the transmission in this case is not subject to the simplifications which have reduced the problems thus far considered to an elementary basis, this matter lies outside the scope of the present paper.[4]

In this development, certain standing waves of sound have been treated on the assumption that the mass and the stiffness in a bounded volume of air could be localized. This idea, while limited in application, works very well in a number of important cases. The characteristic behavior of the single resonator can be explained on this basis. The facts concerning coupled oscillators, which are easily demonstrated in the acoustic case, can thus be handled by students without extensive background. Such work provides an introduction to certain important features in the action of acoustic filters, because it lays a basis for the treatment of the various types of acoustic filters by the same formal analysis used in the case of their electrical and mechanical analogs. Thus, while the subject has interesting implications and very numerous applications, the material which is covered can be adapted to the group for which it is intended.

The writer is indebted to Dr. G. W. Stewart for suggesting the use of the Helmholtz resonator as an experiment for the elementary laboratory and to Mr. W. R. Kennedy for assisting in the construction of the variable frequency sound source used in this work.

[4] For information on this subject, the student may be referred to advanced textbooks on acoustics; for example, Stewart and Lindsay, *Acoustics* (Van Nostrand, 1930), ch. 3 on "Transmission" and ch. 4 on "Filtration of sound."

NOTES AND DISCUSSION

To Prove that the Energy of a Photon (Corpuscle) of Light is Proportional to the Frequency of the Light Wave

GORDON FERRIE HULL
Dartmouth College, Hanover, New Hampshire

IN "A note on the nature of light," Professor Jorgensen[1] presents a proof of the principle that the energy of a light corpuscle is proportional to the frequency of the light wave. In this he assumes that the pressure p of light falling on a totally absorbing body X is equal to the energy density A in the beam. Then he allows X to be pushed against the light with speed u and states that X (of 1 cm² cross section) now receives energy of amount $(c+2u)A$ per second, in place of cA when at rest.

This line of argument has had an interesting history. Drude, in his excellent textbook, *Theory of optics*,[2] derives the relation for the pressure of light, $p=A$, in this fashion. If X be displaced a distance u in the direction of the light (away from the source) the energy which *falls* upon the body is less than before by the amount uA. But the same amount of radiant energy *enters* the body as when at rest. If, therefore, the same energy which *enters* the body develops less heat than when X is at rest, the loss of heat Au must be represented by work done during the displacement. The work is pu. Hence, $p=A$.

Richtmyer[3] and various authors who have followed him avoid Drude's reasoning regarding the difference between *entering* and *falling on* and proceed directly to derive $p=A$ by the simple argument: let X approach the light source with speed u; then more energy Au is received than before; this is due to work done pu and, therefore, $p=A$. In an article entitled "Some fallacies in textbooks on modern physics,"[4] I pointed out that this argument was fallacious. Using the same reasoning we can prove that the pressure is zero. Also, following Richtmyer, we would prove that, for the case of material particles—bullets, for example—striking an absorbing surface, p would equal A, whereas it equals $2A$. So there is danger in basing an argument on the property of absorption. However, Professor Jorgensen[1] is not under obligation to derive the relation for the pressure of radiation. He assumes correctly that $p=A$ for light falling on an absorbing surface. Then he allows the surface to advance against the light. He states that the surface is heated an extra amount because it picks up more energy; it is also heated an *extra amount because work is done*. Now when a piston is pushed against a gas, the heat due to the work done is communicated to the gas, in time to the entire body of the gas if enclosed, and not to the piston. In the case of radiant energy is it communicated only to the radiant energy that enters the body? If so, then the body would be heated on account of work done, and the total excess of heat received by the body would be $2Au$, as Professor Jorgensen has found, and not Au as given by Drude and Richtmyer. If, however, the work done against the radiation is communicated even in part to the radiant energy in front of the body, there would be a question regarding the energy gain in the absorber. The uncertainty becomes greater if the absorber is pushed by the light away from the source. If we follow the gas analogy, the light beam would lose energy on account of work done. But if we allow this loss to appear finally in the absorber, as Professor Jorgensen presumably would do, the absorber, instead of receiving cA, would receive $(c-2u)A$ energy per second. It would be stretching the argument too far to allow u, the speed of the absorber, to equal $\frac{1}{2}c$—then the absorber would not be heated, though radiant energy of amount $\frac{1}{2}cA$ would be entering it per second. (Here, however, the pressure would not equal A.)

The difficulty connected with absorption may be avoided in one of two ways—we may bottle up the radiation in an enclosure with perfectly reflecting walls and allow adiabatic expansion or contraction, or we may direct a parallel beam upon an advancing mirror. In the first method we come to the treatment that I have presented in my textbook.[5] There it is shown that, in an adiabatic compression, λT is constant when T is the absolute temperature of the enclosure. If we look upon light as a corpuscular phenomenon, the number of corpuscles (photons) inside of an enclosure is not changed by an adiabatic expansion. Assume now that the energy E of a photon is proportional to T; then λE is constant and, since $\lambda f=c$, we have $E/f=\text{const.}$

However, the result may be obtained directly by allowing a parallel beam (1 cm² cross section) of radiation to fall upon a plane perfect reflector. Now all energy changes take place in the radiation, none in the reflector. Allow this reflector to advance against the light. Radiant energy of length $c+u$ is reflected to occupy a length $c-u$. The wave-length is changed by the factor $(c-u)/(c+u)$, and the frequency is changed by the inverse factor, or $1+(2u/c)$. The energy $(c+u)A$ has been increased, on account of work done, by an amount $2uA$, since the pressure now is $2A$. Hence, the energy has been increased in the ratio $(c+3u)/(c+u)$, or $1+(2u/c)$. Now the number of photons has not changed and, consequently, the energy of each photon has been increased by the factor $1+(2u/c)$. But that also is the factor of increase for the frequency; hence the result, the energy of a photon is proportional to the frequency, as we use that term in the wave theory of light.

This mode of treatment recalls Larmor's proof that the pressure of radiation is equal to the energy density in the beam.[6] But it was necessary in that argument to assume that the amplitude of the waves was unchanged by reflection from a perfect reflector and that the energy density was inversely proportional to the square of the wave-length. It followed that the energy in the reflected length $c-u$ of the beam was greater than that in the original length $c+u$ by the quantity $2Au$. As this was due to work done pu, then $p=2A$.

It may be noted that the "frequency" of the light entering an absorber cannot be determined merely by measuring the energy received; it must be computed after the "wave-length" has been measured by an appropriate device, a grating, for example. If the light reflected from the mirror is thrown on a grating, a maximum will be found along the line given by $\lambda=a \sin \theta$. If the detector (which must absorb some of the radiation falling on it) is on that line,

it will receive a maximum of energy, whether it is advancing towards or receding from the grating. If it is not on the line determined by the foregoing relation, it will receive no energy. Thus the wave-length, and therefore, the frequency, will be determined by geometrical conditions, not by the characteristics of the detector—except that the latter must detect.

[1] Am. J. Phys. **9**, 243 (1941).
[2] English translation, p. 491.
[3] *Introduction to modern physics*, p. 191.
[4] Am. J. Phys. (Am. Phys. T.) **5**, 22 (1937).
[5] *An elementary survey of modern physics*, p. 419.
[6] Reference 5, p. 57.

Concepts of Potential Difference and Electromotive Force as Presented in College Physics Textbooks

R. W. KENWORTHY
University of Washington, Seattle, Washington

IN the issues of the AMERICAN JOURNAL OF PHYSICS for the past few years one finds some ten articles on the subject of systems of electric units, most of which are concerned with criticism of definitions, concepts, dimensions or uses of the units. Apparently there is some confusion or disagreement in the minds of physicists on these topics. After some years of attempting to teach physics to a general group of university students who have had no previous contact with the field, the writer has been impressed with their confusion of concepts in electricity. The specific important concepts that seem to cause the most difficulty for the student are those of potential difference and electromotive force. A discussion of difficulties with individual students leads one to believe that the most important stumbling blocks are (1) the use of from two to four systems of electric units, (2) a lack of consistency in the definitions of the corresponding units in the different systems. In order to ascertain how generally this criticism could be applied to textbooks, a survey of 22 college physics textbooks published within the last five years was made. The specific topics investigated were:

(1) The definition of potential difference in the electrostatic system;

(2) The definition of the same quantity in the practical system;

(3) The concept of electromotive force.

A definition of potential difference in the electrostatic system was omitted in three books; the remaining 19 defined it as the energy difference in ergs per unit positive charge, which is the generally accepted definition and a most useful one. Consistency of definitions should lead one to expect a similar statement for potential difference in volts. Instead, however, one finds at least five different definitions, such as the following:

The volt is the unit of electrical pressure. (5 books.)
The volt is the potential difference between the ends of a conductor of resistance 1 ohm when the steady current is 1 amp. (3 books.)
The volt is defined in terms of the emf of a standard cell. (2 books.)
The volt is 10^8 emu or 1/300 esu of potential difference. (10 books.)
The volt is the potential difference between two points when one joule of work is required to transfer one coulomb of electricity from one point to the other. (7 books.)

A few textbooks made statements that were inconsistent; for example, in some books the volt was defined both as electric pressure and in terms of the electrostatic or electromagnetic unit of potential difference.

Let us look at these definitions, not from our own point of view, but from that of the student who has had no previous contact with physics. The first definition (that of electric pressure) presumably builds upon his previous experience in the mechanics of fluids; but he soon finds that the analogy offers him only false security, and he must start over again to obtain a useful conceptual idea of potential difference. The second definition (by Ohm's law) is a purely mathematical one, and the student is able to continue farther with it than with the idea of electric pressure; but when he attempts to apply it to calculations of energy changes in motors, generators and storage batteries, he again finds it necessary to seek a new approach. The definition of the volt in terms of the emf of a standard cell gives the student a laboratory standard for comparison purposes only; it has no other real value.

The definition in terms of electrostatic and electromagnetic units is literally correct, but until the corresponding units of quantity and energy are converted to coulombs and joules, it is not a useful definition. The fifth definition —energy change in joules per coulomb—provides the most useful concept for the student, and the seven authors who use it are to be commended. By use of it a student obtains a real grasp of the energy conversions that occur in the various parts of a circuit; $E \pm IR$ for motors, generators and cells becomes a conceptual physical quantity rather than a mathematical manipulation.

Similarly in the case of electromotive force one finds some four different definitions. Three textbooks offered no definite definition of emf. One book defined it by Kirchhoff's second law, another as the maximum electric pressure of a generator. Twelve books defined it as the open circuit potential difference of a cell or generator, and five as the work in joules required to transfer a coulomb of charge completely around the circuit. Which of these is best from the student's viewpoint is somewhat a matter of opinion, but it is obvious that the first three are directly dependent on the correct concept of potential difference. It is the writer's opinion that the last is to be preferred and use of it leads to a real understanding of emf by the student.

The extent of the inaccuracies or disagreements in all of these definitions is not so important as the fact that there *is* disagreement. The most apparent cause of the disagreement is twofold: (1) the use of four systems of electric units; (2) the attempt to substitute analogies and examples for concepts and definitions. Analogies are useful, but almost invariably are carried too far; they should never be used as substitutes for a direct approach. The field of electricity is sufficiently matured so that too much dependence on analogs is unnecessary at this time.

In 1935 the International Electrotechnical Commission decreed that the meter-kilogram-second system should go into effect on January 1, 1940. In 1938 the A.A.P.T. Com-

mittee on Electric and Magnetic Units stated[1] that "the advantages to be gained by fewer systems and more uniformity are so great that it is hoped this report will receive favorable consideration." The advantages are especially important in the case of electric units. All of the elementary textbooks examined were published since 1935, and a very few since 1940, but not one of them makes any use of the mks system.[2] The logical place for introducing a new and simplifying system of units is in the first course in the subject. It is here that the American Association of Physics Teachers should exert a vigorous and forceful leadership. Much has been done in the way of standardization; namely, in methods of teaching geometrical optics, content of secondary school courses in physics, standard tests for college courses, and so forth. Is not here the opportunity for the Association, through its committees and by direct contact with publishers and authors, to bring about as quickly as possible the transition to the new system of units and an immediate strengthening of the foundation work in physics?

[1] Am. J. Phys. (Am. Phys. T.) **6**, 144 (1938), p. 151.
[2] A single elementary textbook that employs the mks system has appeared since this study was made.

An Automatic Control and Timing Device

ALFRED H. WEBER AND EDWARD J. GRILL
Saint Louis University, Saint Louis, Missouri

FIGURES 1 and 2, in which corresponding parts are lettered alike, illustrate a simple device and circuit used to control and automatically time the current in an external circuit. Its features are (1) the control of a relatively large amount of power in an external circuit with a control circuit requiring very little power expenditure for inertial or frictional drags, (2) the timing of the external circuit power for readily varied intervals. We have used the device to "flash" an air-cooled Coolidge x-ray tube for some 5 sec out of every minute and to "rest" the tube by preventing the flashing for some 10 min out of every hour.

The timing of intervals is accomplished with an electric clock, the second hand S and minute hand M of which dip into pools of mercury P in the wooden troughs T_S and T_M (only T_S is clearly visible in Fig. 1; T_M is behind and below), the interval timed being adjusted by varying

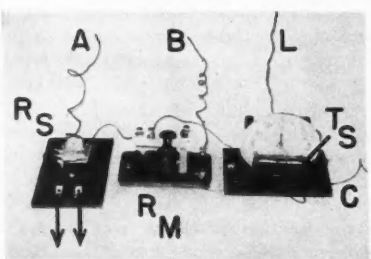


FIG. 1. Photograph of the essential parts of the device: L is the 115-v a.c. line for the clock.

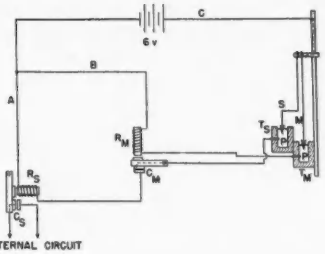


FIG. 2. Circuit diagram.

the amount of mercury in the troughs or the lengths of S and M or both. The hand S is bent as indicated in Fig. 2. From Fig. 2 it is obvious that current is set up in the external circuit whenever S dips into P, thus producing contact at C_S by energizing relay R_S (C_M remaining closed); and the current is interrupted whenever M dips into P, thus breaking contact at C_M by energizing relay R_M. The relay R_S must have a contact C_S designed to carry the heavy current[1] in the external circuit. The relay R_M is an ordinary telegraph sounder.

By the use of multiple-pointer second and minute hands of various lengths, the frequency and duration of the "on" and "off" intervals in the external circuit could be adjusted over a very wide range.

Although the device involves nothing that is very original, we have found it convenient and considered it worth while calling to the attention of others.

[1] The K-100 (Guardian Electric Co., Chicago) relay used will take 1500 w.

An Appeal for Physics Graduates

J. C. MORRIS[1]
National Research Council, Washington, D. C.

THE Office of Scientific Personnel has been set up by the National Research Council at the request of the National Defense Research Committee to assist in locating scientists for the various governmental defense agencies. Working in close cooperation with the National Roster of Scientific and Specialized Personnel it has been active in locating trained scientists in a large number of fields, but by far the greater part of its effort has been devoted to attempting to meet the requests for trained physicists and radio engineers. In both of these fields acute shortages are developing, and there is every indication that the demand will continue to increase.

It is, therefore, imperative that the students who are completing their training in physics during this academic year should be enrolled in the National Roster of Scientific and Specialized Personnel, and if available for defense work should also be listed with the Office of Scientific Personnel. The National Roster plans to obtain in the near future, through the college authorities, the names of those expected to obtain degrees in 1942 in all the fields covered by the Roster. However, because of the pressure that is being felt in the field of physics the National Roster, in cooperation with the Office of Scientific Person-

nel, is preparing to send out at once a special request to the heads of the physics departments asking them to supply the names of those who will receive Bachelor's, Master's or Doctor's degree in *physics* and to indicate which of these will probably be available for defense positions. It is especially urgent that the names of those who will receive their degrees in February should be obtained at once.

The experience of the Office of Scientific Personnel and of the National Roster indicates that the majority of the demands for physicists will be for those with advanced degrees, and *it is hoped that every encouragement will be given to students to continue their training as far as possible.* If the number of those taking advanced training is not maintained the situation will soon become acute. It is also felt that graduate students can play an increased role in assisting in the teaching load, thereby in part compensating for the loss of faculty members who have been called away by defense agencies.

The cooperation of the chairman of every physics department is requested to make this census as complete as possible.

[1] Director, Office of Scientific Personnel; Consultant, National Roster of Scientific and Specialized Personnel.

Frances Gertrude Wick, 1875–1941

THE active and distinguished career of Frances Gertrude Wick was terminated by her death on June 15, 1941. Her illness of about four months duration following a fall suffered in the winter, was born with the fortitude and optimism that characterized all her activities.

Miss Wick was born in Butler, Pennsylvania, on October 2, 1875. Graduating from Wilson College in 1897 she taught for six years in the Butler high school. She then went to Cornell University where she received the degree of bachelor of arts in 1905, the master of arts in 1906 and the doctorate in 1908. After teaching the next two years at Simmons College, she joined the Vassar College department of physics as an instructor in 1910 and was promoted through all the ranks until she became a professor in 1922. At the time of her death she was the chairman of the department, having been elected in 1939.

Professor Wick's outstanding contributions in teaching and research were a consequence of her enthusiasm, friendliness and unfailing energy. She inspired both students and colleagues by her encouragement and by her sincere interest in their work. Always modest concerning her own achievements, she showed a stimulating appreciation of the accomplishments of others. In the classroom her presentations and demonstrations were given each time with the freshness and zeal of the novice but with the skill of the experienced teacher. Her contagious enthusiasm was communicated to her students, many of whom became permanently interested in the field of phsyics.

In summer vacations as well as during various leaves Professor Wick worked in many physical laboratories, among them that of Cornell University, the Jefferson Physical Laboratory of Harvard University, the Research Laboratory of the General Electric Company, the Cavendish Laboratory as the guest of J. J. Thompson, and the Institute for Radium Research of Vienna. Her researches were chiefly in the field of luminescence, in which she was in turn the student of, collaborator with and successor to the late Professor E. L. Nichols of Cornell University. Many of her investigations were pioneer studies carried out before phenomena of luminescence had assumed the importance that they now have. In spite of ailing health Miss Wick had continued her work and was anticipating with great pleasure an increasing use of the extensive and unique collection of luminescent materials which had been prepared by Professor Nichols and placed at her disposal by Cornell University.

It was not alone in teaching and research that Miss Wick was active. Her broad interests were indicated by extensive travels, during which she enjoyed the life of the people of various countries and became interested in many scientists, some of whom she was able to help in their misfortunes. She also served as a trustee of Wilson College and on committees of the National Research Council, the American Association of University Women and the Optical Society of America.

In the minds of all who knew her Miss Wick will live as a remarkable example of contagious happiness and constant devotion to her chosen work.

PAUL A. NORTHROP

"FOR I dipt into the future, far as human eyes could see,
Saw the vision of the world, and all the wonders that would be,
Saw the heavens fill with commerce, argosies of magic sails,
Pilots of the purple twilight, dropping down with costly bales."

ALFRED TENNYSON (1840)

RECENT PUBLICATIONS AND TEACHING AIDS

Survey Courses

Fundamentals of Physical Science. Konrad Bates Krauskopf, Assistant Professor of Geology, Stanford University. 670 p., 329 fig., 15×23 cm. *McGraw-Hill*, $3.50. Professor Krauskopf has presented to the teaching profession an outstanding example of what can be done in the field of science when one views the subjects of physics, chemistry, astronomy and geology as a whole. The book should be excellent for the student who is really interested in the field of science. It is not of the usual "elementary" type but is comparable in its treatment with any standard college scientific textbook.

A reasonable amount of mathematics is used. The definitions are precise, and the explanations, while in some cases extensive, are interesting and exact. An excellent set of questions appears at the end of each chapter, and a list of reference material at the end of each section. The book is divided into six parts for convenience in using it under either the semester or term system; but the material is fairly well distributed in the four fields and is integrated to the extent that the treatment is consecutive and logical. Although this treatment is quite thorough, it is not considered too difficult for the average college freshman.

The approach is largely historical, with particular reference to the scientific experimental method. Excellent examples of good presentation are the discussions of inertia, mass and weight, and the chapter on mathematics. The book contains a liberal number of chemical equations and formulas but does not introduce the more difficult ones.

This textbook would be most satisfactory for a course in physical science survey covering the four fields indicated, where the intent of the course is to teach the student to think in terms of scientific procedure rather than to give him descriptive knowledge of scientific material.—Will V. Norris.

This Physical World. C. C. Clark and C. A. Johnson, New York University, and L. M. Cockaday, United States Naval Academy. 538 p., 15×23 cm. *McGraw-Hill*, $3.25. The authors state in the preface that their aim is to present in a readable style an accurate discussion of some of our basic knowledge about the physical world. The book is intended as a freshman survey textbook on the college level covering the fields of astronomy, physics and chemistry. It opens with the historical background of Tycho Brahe and ends with a chapter on communications. Primarily about science, rather than a study in science, the book contains no laboratory material and is nonmathematical; but it does give an interesting discussion of the experimental fields covered. Sets of questions for students' study are missing. The references for additional reading are particularly good in that they include a brief description of the contents of each reference. Unusual emphasis is placed on the fields of electricity and electrical engineering. The material on astronomy is good, as is the general description of atomic structure. Chemistry is well covered, particularly in reference to modern synthetic products. The diagrams are excellent and instructive, particularly those relating the British thermal unit to the foot pound, the one on heat quantities and the diagrammatical ladder of dimensions. Some of the photographs appear to be unnecessary; for example, pouring gold, the oxygen tent, the Lucite retort and the star diagram. The type is large and readable.

Although the explanations are for the most part quite elementary, they are thorough and illustrate the use of the scientific method. The technical definitions used might be questioned in a few cases. However, the discussion is quite satisfactory for students in a descriptive course who wish to learn some of the interesting facts about science, presented under such chapter headings as "Electrons in glass houses."—Will V. Norris.

Meteorology

Introduction to Meteorology. Sverre Petterssen, Professor of Meteorology, Massachusetts Institute of Technology. 245 p., 143 fig. and plates, 15×23 cm. *McGraw-Hill*, $2.50. Designed to serve as an elementary introduction to modern meteorology, this textbook is intended to create an interest in the subject and provide a background rather than to furnish a technical treatment. The book is an expansion of a chapter on meteorology written for the British and American editions of Weem's *Air navigation* and is an abbreviation of the general principles of the author's recent book, *Weather analysis and forecasting*.

After a very brief chapter devoted to the composition and structure of the atmosphere, the author discusses instruments and observations, then evaporation, condensation and precipitation. The effects of temperature changes in the atmosphere, both adiabatic and nonadiabatic, are very adequately treated, especially the adiabatic changes and the stability and instability of the air. The subject of wind systems is treated in a fashion that seems to be a little more than elementary. A somewhat brief chapter on air masses is followed by a quite thorough discussion of the formation, structure and behavior of fronts; the treatment of fronto-genesis is excellent. The growth of cyclones and anticyclones is fully discussed.

The chapters devoted to weather analysis, weather forecasting and weather maps are perhaps the best parts of the book. The methods followed and the symbols used in constructing weather charts are given in detail; the steps followed in making a complete analysis of the charts are described. The methods used in forecasting are given in detail and with clarity by the aid of some excellent weather maps. The book is concluded with brief chapters on climate and the history of meteorology.

The arrangement of some of the subject matter might be improved and some parts seem rather advanced for the elementary and brief introductions preceding them. On the whole, however, the book is attractive and well written and doubtless will fill well the purpose for which it was written.—John G. Albright.

Astronomy

Astronomy. Clyde Fisher and Marian Lockwood, Hayden Planetarium, American Museum of Natural History. 214 p., 66 fig., 14×22 cm. *Wiley*, $1.75. This is a well-organized and well-written compendium of facts concerning the universe. The material has been carefully chosen and is in good balance. The relative amount of space given each topic in general meets with the approval of this reviewer. The several theories which have been chosen for delineation are described satisfactorily.

The book should prove to be of interest to the superficial reader. However, it falls short of what the editor of this series believes should be the aim of the several books in the series. He says, "It is not difficult to sketch the scene superficially but the result is a smattering of descriptive knowledge which is far from being a science. . . . The student and the citizen need to absorb the scientific attitude, to master the scientific method of thought and to master the basic concepts of the science." With all of this the present reviewer is enthusiastically in accord. The book does little more than sketch the scene. In certain places the wording suggests that the authors themselves are not thinking very clearly; for example, on page 47 they say, "Obviously, since it is not a solid but a gaseous body, the sun need not and does not rotate at its various latitudes with a uniform speed." It seems like a truism to the writer that a book written with the ideals stated by the editor should show *how* ideas grow from observations. Here, and in countless other places, the book fails to do this when the opportunities are ripe.

This reviewer has yet another criticism. In his experience with students in the college freshman group, he finds this coterie of inquiring minds to be as much if not more interested in *how* the scientist arrives at a conclusion as they are in the conclusion itself. A typical example concerns the temperature of the moon. The instructor says hopefully, "The temperature of the moon is above the boiling point of water when the sun is directly over head." More students immediately ask, "How in the world can you tell?" than the simpler question "Why?". Obviously, the second question is naturally included in the answer to the first. Furthermore, if the first question is properly answered, the scientific method can be presented with meaning. This the present book fails to do. To be sure, there is a chapter on instruments—considered separately, as are the planets, meteorites, and so forth—but the material is not handled as well as is the factual matter concerning the universe. For instance, concerning the photoelectric cell the book says "The photoelectric cell developed by Stebbins and Rosing has made it possible to measure with great accuracy the magnitudes of stars, a very important advance in the study of stars." Here is a weak answer to the question, "What is it good for?", and no answer at all to the more frequently asked questions, "What is it like?" and "How is it used?". The references to the development and great accuracy of the photoelectric cell are also obviously misleading and indefinite.

In conclusion, the reviewer regards this book as above the average in clarity, simplicity of description and balance in choice of material. It fails, however, to meet the ideals expressed by the editor of the series and in so doing fails to meet a requirement of any good scientific book, including those intended for the unscientifically trained reader.—C. E. Hesthal.

Motion Picture Films

Unseen Worlds. 16 or 35 mm, sound, 10 min. *Wm. J. Ganz Co.* (19 E. 47th St., New York), loaned gratis. Electron microscope.

Slidefilms

Education Slidefilms. 35 mm. *Jim Handy Organization* (2821 East Grand Blvd., Detroit), sale only. Twelve sets of film slides on magnetism, static electricity, current electricity, primary cells, secondary cells, electromagnetism, generators, alternating current, motors, meters, applications.

National Roster of Scientific and Specialized Personnel

THE American Institute of Physics wishes to emphasize the fact that every American physicist should be registered in the National Roster of Scientific and Specialized Personnel. This is true for holders of Bachelor's degrees as well as holders of more advanced degrees. Your registration will help defense authorities to estimate the scientific capacity of the nation and to plan projects and training programs accordingly. It may help to assure you appropriate opportunities for work in National Defense and may save you from assignments in which your specialized ability would be wasted.

If you *have not* enrolled, send at once for a questionary and a check list for your field. Be sure to specify the field—physics, radio, chemistry or whatever it is. If you *have* enrolled, you will shortly receive a new questionary designed to bring up to date the information in the Roster; do not fail to return it completely answered.

It is especially important that all physicists already engaged in National Defense work be correctly registered in the Roster; this will forestall recommendations for personnel transfers that would be disturbing both to the individuals concerned and to defense projects and industries. It is also essential for the Government to know what percentage of the experts in any field are employed on defense projects and thus to anticipate serious shortages and deal with them intelligently.

Physics departments are urged to canvass their staffs and graduate students and to submit the names of any not registered to the National Roster of Scientific and Specialized Personnel, 930 F Street, N. W., Washington, D. C.

Author Index to Volume 9

In this index are listed the names of the authors and the titles of their articles. Abstracts of papers read at meetings and articles that appear in the "Digest of Periodical Literature" section of the journal are not listed; they are indexed in the Analytic Subject Index.

Analytic Subject Index to Volume 9

The titles of articles are disregarded, the entries being based on analyses of the contents of the original articles. Entries marked (D) refer to digests appearing under "Digests of Periodical Literature" and to abstracts of papers read at meetings; entries marked (R) refer to reviews appearing under "Recent Publications and Teaching Aids."

To facilitate reference to any desired subject, the index is divided into sections arranged alphabetically. The titles of these sections and of various cross-references to them are as follows: